BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

Patent Office
Canberra

AU 00/00864

I, KAY WARD, TEAM LEADER EXAMINATION SUPPORT AND SALES hereby certify that annexed is a true copy of the Provisional specification in connection with Application No. PQ1675 for a patent by ST. VINCENT'S INSTITUTE OF MEDICAL RESEARCH filed on 19 July 1999.

EJKW

WITNESS my hand this
Twenty-sixth day of July 2000

KWard

KAY WARD
TEAM LEADER EXAMINATION
SUPPORT AND SALES

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

AUSTRALIA

Patents Act 1990

PROVISIONAL SPECIFICATION

Applicant(s):

ST. VINCENT'S INSTITUTE OF MEDICAL RESEARCH

Invention Title:

INHIBITOR OF OSTEOCLAST PRECURSOR FORMATION

The invention is described in the following statement:

INHIBITOR OF OSTEOCLAST PRECURSOR FORMATION

This invention relates to a polypeptide factor which is able to inhibit the formation of osteoclasts. In particular, the invention relates to a factor which inhibits the differentiation of haematopoietic precursor cells into mononucleate osteoclast precursors. In a preferred form of the invention, the factor is a type II membrane polypeptide expressed on the osteoblast cell surface, which we have designated osteoclast inhibitory lectin (OCIL).

BACKGROUND OF THE INVENTION

In normal adults, the processes of bone formation and resorption are balanced in order to maintain a normal healthy bone mass. With the onset of the menopause in females and with ageing in both sexes, the rate of bone resorption exceeds that of bone formation, resulting in net bone loss, and ultimately in osteoporosis.

Osteoblasts are the bone cells responsible for bone formation, while osteoclasts are responsible for resorption of bone. Our understanding of the factors that regulate the formation and function of osteoclasts has been greatly enhanced by laboratory methods that have enabled us to isolate and grow these cells in culture. It is now well established that the development of active osteoclasts *in vitro* requires intimate contact between osteoblastic stromal cells and precursors of osteoclasts which are derived from haematopoietic cells belonging to the monocyte/macrophage lineage (Takahashi *et al*, 1988). This process is influenced by a variety of factors, including 1,25-dihydroxyvitamin D_3, parathyroid hormone, prostaglandin E_2, and interleukins 6, 11 and 17, all of which enhance osteoclast formation. In contrast, cytokines such as interleukins 4, 10, 13 and 18 are inhibitory (Suda *et al*, 1995; Martin and Udagawa, 1998).

All factors which stimulate osteoclast formation act directly on the osteoblast population and not on the osteoclast precursors, leading to the proposal that osteoblasts or stromal cells express a membrane-associated peptide that regulates the formation of functional multinucleate osteoclasts. A factor, termed "Osteoclast Differentiation Factor" (ODF), that fulfils the functions of such a putative membrane-associated peptide has recently been cloned. ODF encodes a 316 amino acid type II transmembrane protein, and is a member of the TNF ligand family (Yasuda *et al*, 1998). Recombinant protein corresponding to the extracellular domain of ODF stimulates the formation of active, bone-resorbing osteoclasts from haematopoietic cells within the spleen, even in the absence of stromal cells. A peptide identical to ODF has also been cloned from T cells, and designated Tumour Necrosis Factor-related activation-induced cytokine (TRANCE), or receptor activator of NF-κB ligand (RANKL) (Wong *et al*, 1997). When released by T cells following activation of the T cell receptor, it mediates the interaction of T cells and dendritic cells, resulting in stimulation as well as increased survival of the naïve T cells. RANK, another member of the TNF-receptor family, has been identified on dendritic cells, and acts as the receptor for TRANCE/RANKL (Anderson *et al*, 1997).

Osteoprotegerin (OPG) is a soluble factor that belongs to the Tumour Necrosis Factor (TNF) receptor family. This factor is also known as Osteoclastogenesis Inhibitory Factor (OCIF). It has been shown to bind to ODF, resulting in the inhibition of formation of functional multinucleate osteoclasts *in vitro*. OPG is a 401 amino acid polypeptide. Overexpression of OPG in transgenic mice results in severe osteopetrosis, with a loss of bone marrow cavities and profound depletion of osteoclasts. The same effects were observed upon administration of OPG to normal mice. Furthermore, OPG blocked ovariectomy-associated bone loss in rat. OPG mRNA transcripts have been identified

within bone and cartilage, vascular structures, midgut and kidney, and in several osteoblast cell lines. Current data suggest that OPG blocks the terminal stages of osteoclast differentiation, but not the formation of mononuclear osteoclast precursors (Simonet *et al*, 1997; Tsuda *et al*, 1997).

The interaction between ODF and OPG in the formation of osteoclasts is illustrated in Figure 1. Osteoclasts are derived from haematopoietic stem cells that differentiate along the monocyte/macrophage lineage. Mononuclear precursors of osteoclasts are required to come into direct or close contact with osteoblasts to be rendered capable of differentiating into mature, functional, multinucleate osteoclasts. Osteoblasts express ODF, a membrane-bound protein that stimulates the differentiation and formation of multinucleate osteoclasts from mononuclear precursors when it binds to its receptor, RANK. ODF expression is stimulated by bone-resorbing factors such as PTH, PGE_2, 1,25-dihydroxyvitamin D_3 and interleukins 6 and 11. The action of ODF is antagonised by Osteoprotegerin, a soluble factor secreted by osteoblastic stromal cells. It binds to ODF to inhibit the formation of differentiated multinucleate osteoclasts, but does not prevent the formation of mononuclear osteoclast precursors.

We have now identified a polypeptide factor which is able to inhibit formation of mononuclear osteoclast precursors from haematopoietic stem cells, and which is expressed at least on the cell membranes of osteoblasts. It appears that when the molecule is expressed on the osteoblast cell membrane it is not secreted. Preventing expression of the factor results in increased formation of mononuclear precursors of osteoclasts.

SUMMARY OF THE INVENTION

In a first aspect, the invention provides a nucleic acid molecule which comprises a sequence encoding a protein which

a) is expressed at least on osteoblasts, and

b) inhibits osteoclast differentiation from haematopoietic cell precursors,

or which hybridises to said nucleic acid molecule under stringent conditions. Suitable stringent conditions are well known in the art. See the well known textbook by Sambrook *et al* (1989),and Example 2 herein.

The nucleic acid may be cDNA, genomic DNA or messenger RNA. Preferably the nucleic acid molecule is a cDNA. More preferably the cDNA comprises a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10 and SEQ ID NO: 11.

Preferably the protein inhibits differentiation of haematopoietic stem cells to osteoclast progenitor cells. In a particularly preferred embodiment, the nucleic acid molecule of the invention comprises a 110 base pair sequence as set out in SEQ ID NO: 2.

This aspect of the invention also encompasses anti-sense sequences directed against the nucleic acid molecule defined above, and particularly encompasses an anti-sense sequence directed against SEQ ID NO: 10. Preferably the anti-sense sequence is SEQ ID NO: 24 or SEQ ID NO: 25.

In a second aspect, the invention provides a polypeptide encoded by the nucleic acid molecule of the invention. Preferably the polypeptide is encoded by the human cDNA sequence. More preferably the polypeptide comprises an amino acid sequence encoded by SEQ ID NO: 20.

In a third aspect, the invention provides an antibody directed against a polypeptide of the invention. Preferably the antibody is directed against an epitope present in the sequence

H-Cys-Met-Arg-Thr-Glu-Ala-Gln-Leu-Ala-Arg-Phe-Asp-Asn-Gln-Asp-Glu-Leu-Asn-Phe-OH (SEQ ID NO: 26).

The antibody may be polyclonal or monoclonal, but is preferably monoclonal. Suitable methods for generating either polyclonal or monoclonal antibodies are very well known in the art. It will be clearly understood that the invention encompasses biologically-active fragments and analogues of such antibodies, including but not limited to ScFv fragments, trimeric antibodies, humanised antibodies and the like. Again, methods for producing such active fragments and analogues are well known in the art. See for example PCT/AU93/00491 and PCT/AU97/00212 and references cited therein.

In a fourth aspect, the invention provides a composition comprising a polypeptide or an antibody of the invention, together with a pharmaceutically-acceptable carrier.

Methods and pharmaceutical carriers for preparation of pharmaceutical compositions are well known in the art, for example as set out in textbooks such as Remington's Pharmaceutical Sciences, 17th Edition, Mack Publishing Company, Easton, Pennsylvania, USA.

In a fifth aspect, the invention provides a method of treatment of a condition characterised by abnormal bone resorption, comprising the step of administering an effective amount of a modulator of expression or function of the polypeptide of the invention.

Where the condition involves excessive bone resorption, the method will comprise administration of the polypeptide of the invention or the nucleic acid encoding this polypeptide, or a biologically-active fragment or analogue thereof. Such conditions include, but are not limited to, osteoporosis, primary hyperparathyrodism, Paget's disease, rheumatoid arthritis, renal osteodystrophy, humoral hypercalcaemia of malignancy, and conditions where cancer has metastasised to bone.

Conditions characterised by deficient bone resorption include osteopetrosis. Antibodies directed against the polypeptide of the invention or anti-sense

oligonucleotides directed against the nucleic acid of the invention may be used to inhibit the function of the polypeptide and thus to increase bone resorption.

It is also contemplated that the polypeptide of the invention may be used to promote healing of bone fractures, particularly in individuals where fracture healing is delayed or deficient. These include individuals suffering from osteoporosis or diabetes mellitus.

In a sixth aspect, the invention provides a diagnostic kit, comprising a reagent selected from the group consisting of a nucleic acid of the invention or a fragment thereof capable of hybridising to a nucleic acid of the invention; an anti-sense nucleic acid of the invention; a polypeptide of the invention, and an antibody of the invention. For example, diagnostic kits for use in methods such as polymerase chain reaction, fluorescent *in situ* hybridisation, immunoassay, and the like are contemplated. Where appropriate, the molecule of the invention which is used may be labelled with a detectable marker, such as a radioactive, fluorescent, chemoluminescent or enzymic marker. Such diagnostic kits are useful for detection of abnormalities in the structure, expression or control of the factor of the invention, which may lead to increased bone resorption and concomitant pathological manifestations.

For the purposes of this specification it will be clearly understood that the word "comprising" means "including but not limited to", and that the word "comprises" has a corresponding meaning.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 summarises the factors and mechanisms involved in control of osteoclast differentiation and development, as understood before the date of the present invention.

Figure 2 shows detection by Northern blotting of a 780 base pair mRNA species using rOCIL323 as a probe in a variety of rat clonal osteoblast-like cell lines.

Figure 3 shows the results of Northern blot analysis using rOCIL402, a 402 base pair fragment obtained by screening of a rat ROS 17/2.8 cDNA library, using the polymerase chain reaction. Similar results were obtained using rOCIL323 fragment as a probe.

Figure 4 shows the comparison of rOCIL323 and rOCIL402 probes in Northern blotting of mRNA from 1,25-dihydroxyvitamin D_3-treated rat UMR 106-06 cells. The results showed that both fragments detected the same species of mRNA.

Figure 5 shows the results of Northern blot analysis of UMR 201 mRNA using mOCIL 2kb, showing that this probe detected the same 780 bp species as rOCIL323 and rOCIL402.

Figure 6 shows the intron-exon structure of the mouse OCIL gene

Figure 7 shows sequence comparisons between human, rat and mouse OCIL, as generated using the program Clustal W. Regions of sequence identity are shown shaded and in bold.

Figure 8 shows comparisons between the deduced protein sequence corresponding to hOCIL and the deduced protein sequences corresponding to rOCIL 1.3 and mOCIL, as generated using the program Clustal W.

Figure 9 shows the results of treatment of cocultures of primary mouse calvarial osteoblasts and mouse bone marrow cells with an anti-sense oligonucleotide, 323 A/S, directed against the C-type lectin region of OCIL and a second antisense oligonucleotide, 474 A/S, which was directed against the sequence in the open reading frame but outside the C-type lectin region.

a: cocultures treated with anti-sense oligonucleotide under basal conditions.

b-d: cocultures stimulated with anti-sense oligonucleotides in the presence of 1,25-dihydroxyvitamin D_3 and PGE_2.

Figure 10 shows the results of Northern blot analysis of mRNA from UMR106 parental cells, demonstrating upregulation of expression of OCIL by IL-1α, IL-1β, IL-11, IL-17, TNFα, TGFβ, M-CSF, GM-CSF, PGE_2, 1,25-dihydroxy-vitamin D_3, 1,25-dihydroxyvitamin D_3 plus PGE_2, and PGE_2 plus dexamethasone.

Figure 11A shows the results of a time-course study, showing upregulation of OCIL by PTHrP.

Figure 11B shows that the upregulation could be detected using either rOCIL402 or mOCIL 2kb as the probe.

Figure 12 shows upregulation of expression of OCIL in primary mouse calvarial osteoblasts by IL-1α, IL-1β, IL-11, dexamethasone, and 1,25-dihydroxyvitamin D_3.

Figure 13A shows upregulation of expression of OCIL in ST2 mouse stromal cells by PGE_2, dexamethasone, 1,25-dihydroxyvitamin D_3, IL-11, PTH, and 1,25-dihydroxy-vitamin D_3 plus PGE_2.

Figure 13B shows the time course of upregulation of OCIL expression in ST2 cells by dexamethasone.

Figure 14 shows the results of Northern blot analysis of adult mouse tissues (left panel) and adult rat tissues (right panel), demonstrating expression of OCIL.

Figure 15 shows the effect of recombinant rOCIL protein on formation of multinucleate osteoclasts from mouse calvarial osteoblasts.

DETAILED DESCRIPTION OF THE INVENTION

The invention will now be described in detail by way of example only, with reference to the following non-limiting examples and drawings.

We set out to clone a gene encoding a peptide that would function to prevent osteoclast formation. It is known that mature osteoblasts have limited potential to support osteoclast formation, and we postulated that mature

osteoblasts might express osteoclastogenic inhibitors. The pre-osteoblastic cell line UMR201 can be differentiated to a more mature osteoblastic phenotype by treatment with 10^{-6} M retinoic acid for 24 hr (Ng *et al*, 1988). mRNA species differentially expressed between mature osteoblasts (retinoic acid-treated UMR201 cells) and immature osteoblasts (untreated UMR201 cells) were identified using an array of oligonucleotide primers in reverse transcription PCR, where products amplified from RNA from the two cellular populations were compared. We characterised products which were elevated in mature osteoblasts as candidates for osteoclastogenic inhibitory molecules.

Abbreviations used herein are as follows:

GM-CSF	granulocyte/macrophage colony stimulating factor
hPTH	human parathyroid hormone
IGF	insulin like growth factor
IL	interleukin
LIF	leukaemia inhibitory factor
M-CSF	macrophage colony stimulating factor (CSF-1)
$1,25(OH)_2D_3$	1,25-dihydroxyvitamin D_3
OCIF	osteoclastogenesis inhibitory factor
ODF	osteoclast differentiation factor
OPG	osteoprotegerin
PCR	polymerase chain reaction
PGE_2	prostaglandin E_2
PTH	parathyroid hormone
PTHrP	parathyroid hormone-related protein
RANK	Receptor activator of NF-κB
TGF	transforming growth factor
TNF	tumour necrosis factor
TRAP	tartrate-resistant acid phosphatase.

Example 1 Isolation of Rat cDNA Encoding the Inhibitory Factor

Total RNA was isolated from retinoic acid-treated preosteoblastic UMR201 cells using guanidine thiocyanate (Chomczynski *et al*, 1987). First strand cDNA was synthesised from 2 µg of total RNA by incubating for 1 h at 42°C with 15 units of AMV reverse transcriptase (Promega, Madison, WI) following oligo dT priming. A sense primer that was complementary to rat calcitonin cDNA, designated primer CT1 (SEQ ID NO:1),

CT1 5'-ATG CTG GGC ACG TAC ACA CAA-3'

and 3'UAP 5'-CUA CUA CUA CUA GGC CAC GCG TCG ACT AGT AC-3' (Clontech) were used as primers in polymerase chain reaction (PCR). The PCR conditions utilised a touchup PCR protocol with denaturation at 94°C for 5 min, and then 5 cycles at 94°C for 1 min, 37°C for 1 min and 72°C for 1 min, followed by 35 cycles of 94°C 1 min, 49°C for 1 min and 72°C for 1 min. For these experiments, Expand High Fidelity PCR System (Boehringer Mannheim) was used in a Perkin Elmer Cetus 480 thermal cycler. A 321 bp PCR product was obtained. This 321 bp fragment, which we designated rOCIL323 (SEQ ID NO: 2), was used as a probe in Northern blots. As shown in Figure 2, it hybridised to a 780 bp mRNA species in UMR 201, UMR 201-10B, UMR 106-06, UMR 106-01 and ROS 17/2.8 cells, all of which are rat clonal osteoblast-like cell lines.

Since retinoic acid and PTH enhance OCIL mRNA expression dramatically in UMR 106-06 and UMR 106 parental cells, a similar RT-PCR procedure was carried out using RNA isolated from retinoic acid or hPTH 1-34 treated UMR 106 parental cells. A PCR product identical to the 321 bp fragment for rOCIL323 was obtained and its expression was found to be upregulated in UMR 106 cells treated with either retinoic acid or hPTH 1-34.

To extend the sequence of OCIL, anchored PCR was used to screen a rat ROS 17/2.8 cDNA library with λgt11

arms. An antisense 21 bp primer, designated OCILr1 (SEQ ID NO: 3),

OCILr1 5'-TGA GTG TTG TCT GTC CAC TTC CAA G-3'

complementary to a sequence in the 321 bp fragment, was used with either the λgt11 forward primer 5'-GGT GGC GAC GAC TCC TGG AGC C-3' or λgt11 reverse primer 5'-GAC ACC AGA CCA ACT GGT AAT G-3' (Clontech) to amplify an aliquot (10^6 plaque forming units) of the recombinant library. Cycling parameters were 94°C for 5 min, then 80 cycles of 94°C for 30 sec, 60°C for 30 sec, and 72°C for 2 min, followed by a final extension step of 72°C for 10 min. A 402 bp fragment was obtained with λgt11 reverse primer as the anchored primer. Sequencing of this 402 bp fragment showed 88.6% identity over a length of 97 bp with rOCIL323 (SEQ ID NO: 2). The 402 bp fragment, designated rOCIL402, whose sequence is set out in SEQ ID NO: 4, was used to probe Northern blots obtained from the rat osteoblast-like cell line. It hybridised to the same 780 bp mRNA species observed with the rOCIL323 probe. These results are shown in Figures 3 and 4. The same results were obtained in both the presence and absence of 1,25-dihydroxyvitamin D_3.

A 3' Rapid Amplification of cDNA Ends (3'-RACE) strategy was used to obtain the 3' ends of the cDNA of interest. The sense specific primers used were OCILr11 (SEQ ID NO: 5)

OCILr11 5'-GAA ACA TCC CCC TGG AGT ATC C-3'

and OCILr12 (SEQ ID NO: 6)

OCILr12 5'-CCA AGT AAC TGG ACA TTG AGC CAG A-3'

complementary to sequences within rOCIL402 (SEQ ID NO: 4). First-strand cDNA was synthesised from total RNA isolated

from hPTH 1-34 treated UMR 106 parental cells, using the oligo dT-anchor primer. The cDNA was further amplified by PCR using OCILr11 or OCILr12 and the 3' adaptor primer AP-1 (Clontech). PCRs were run at 94°C for 5 min, then 35 cycles of 94°C for 30 s, 62°C for 30s, and 72°C for 2 min, followed by a final extension step of 72°C for 10 min. Three different polyadenylated 3' sequences were obtained, designated rOCIL1.3kb (SEQ ID NO: 7), rOCIL738bp (SEQ ID NO: 8) and rOCIL620bp (SEQ ID NO: 9) respectively. The region of sequence identity between rOCIL323 and rOCIL402 was found to extend to 117 bp.

Example 2 Isolation of Mouse cDNA and gDNA Encoding the Inhibitory Factor

rOCIL402 was labelled with [^{32}P] α-dCTP by using the Random Primer labelling kit (Boehringer Mannheim), and a mouse liver cDNA library was subjected to hybridisation screening at 65°C in a hybridisation buffer containing 4 x SSPE (SSPE contains 0.15 M NaCl, 0.01 M NaH_2PO_4, and 0.001 M EDTA), 5 x Denhardt's solution, 0.5% sodium dodecyl sulfate (SDS) for 24 hr. The filters were then washed sequentially in 2 x SSC at 65°C for 15 min, 2 x SSC with 0.1% SDS at 65°C for 30 min, and finally 0.1 x SSC at 65°C for 10min. We obtained a 1907bp mouse cDNA, designated mOCIL2kb (SEQ ID NO: 10). The sequence of mOCIL2kb shows 80% identity over a length of 461 bp to that of rOCIL1.3kb. When used as an antisense riboprobe in Northern blot analysis, mOCIL2kb hybridised to a 780 bp mRNA species in UMR201 as detected by rOCIL323 and rOCIL402,as shown in Figure 5.

A cDNA fragment corresponding to the nucleotides 58-776 of mOCIL2kb was used as a cDNA probe to screen a genomic BAC Mouse I Hybridisation library. The screening α was performed under contract by Genome Systems, Inc. According to their protocol, the cDNA fragment was labelled with [^{32}P] α-dCTP by random primer labelling and the library was screened under the hybridisation condition of

55°C in a hybridisation buffer containing 5.5 x SSC, 5 x Denhardt's solution, 0.5% SDS and 0.5 x HEPES buffer for 18 hr. The filters were then washed sequentially in 1mM Tris-HCl pH 8.0 and 1% sarkosyl for 15 min, and 3 times in 1mM Tris-HCl pH 8.0 for 15 min washes at room temperature. Eight positive clones were isolated, of which four positive clones were screened. After the genomic DNA was digested with HindIII or BamHI, Southern blot analyses were carried on with the same cDNA probe. The results showed that two clones were identical, and one clone could not grow in culture. One of the duplicate clones, designated db.20149, was sequenced by subcloning into the pBS vector and by direct sequencing of the genomic clone using cycle sequencing. Sequencing of 9349 bp of the genomic clone of mOCIL was completed (SEQ ID NO: 11). It contains 6 exons, as shown in Figure 6. The 5' flanking region adjacent to exon I is a GC-rich region, containing a Sp 1 binding site. A 1 kb gap between nucleotides 1862 and 1868 is currently being sequenced. The genomic sequence confirmed that nucleotides 1-320 of the original mOCIL2kb represented an inverted repeat of the 3' end of the sequence, and that nucleotides 900-1907 of mOCIL2kb were not part of the mOCIL sequence. The corrected full length mOCIL cDNA is 990 bp in length (SEQ ID NO: 12).

To confirm the mOCIL cDNA sequence, RT-PCR was carried out using total RNA isolated from ST2 mouse stromal cells, primary mouse calvarial osteoblasts and mouse liver tissue. The sense primer represented nucleotides 18-36 of mOCIL and is designated as primer OCILm47 (SEQ ID NO: 13),

OCILm47 5'- TCC CAT GCC AGA TTG CTT G-3'

The antisense primer represented nucleotides 743-720 of mOCIL and is designated primer OCILm12 (SEQ ID NO: 14),

OCILm12 5'-GGG ACC ATA GGG GAA AAA GTA G-3'

The PCR was run at 94°C for 5 min, then 35 cycles of 94°C for 30s, 60°C for 30s, and 72°C for 1 min, followed by a final extension step of 72°C for 10 min. Five clones containing a 721 bp fragment were obtained from three sources, ST2 mouse stromal cells, primary mouse calvarial osteoblasts and mouse liver tissue. When compared to the mOCIL sequence there was 100% identity in the first 106 bp (exons I and II) but only 90.5% identity in the remaining 615 bp. This 721 bp fragment was designated as mOCIL47 (SEQ ID NO: 15).

To confirm whether nucleotides 320-434 of mOCIL2kb are part of mOCIL, a sense primer representing nucleotides 343-364 of mOCIL2kb, designated as primer OCILm17 (SEQ ID NO: 16),

OCILm17 5'-TGG AAA CTC AGC TCC TCA GCT CTG-3'

and antisense primer OCILm12 was used to carry out RT-PCR with RNA from three sources, ST2 mouse stromal cells, primary mouse calvarial osteoblasts and mouse liver tissue, as above. PCR was run under the same conditions as for mOCIL47. Twelve clones were obtained, each containing a 713 bp fragment. The first 115 bp were not identical to mOCIL. In 2 of the 12 clones, there was 100% identity to mOCIL after the first 115 bp. In the other 10 clones, there was 92.2% identity to mOCIL and 89.2% identity to mOCIL47 after the first 115 bp. This sequence is designated mOCIL17 (SEQ ID NO: 17).

RT-PCR was also carried out using a sense primer corresponding to the region located at the junction of exon III and exon IV, representing nucleotides 243-267 of mOCIL, and designated primer OCILm32 (SEQ ID NO: 18),

OCILm32 5'- TTT GTC AGC AAC AAA GAC AGA ACA G-3'

Primer OCILm12 was used as a antisense primer. RT-PCR was carried out with RNA from three sources, ST2 mouse stromal cells, primary mouse calvarial osteoblasts and mouse liver tissue, as above. PCR was run under the same conditions. Four clones were obtained, each containing a 502 bp fragment. Three of the four clones have 100% identity to mOCIL17 and one of the four clones is 100% identical to mOCIL.

mOCIL has an open reading frame encoding a 218 amino acid protein. Its structure is typical of a type II membrane protein, with a predicted 142 amino acid extracellular domain, a 21 amino acid transmembrane domain, and a 55 amino acid cytoplasmic domain. The extracellular domain has 6 cysteine residues. There are three potential N-linked glycosylation sites at residues 86, 112 and 207, all of which are in the extracellular domain. There is a myristylation motif in the intracellular domain.

Comparison of the putative protein sequences derived from the rOCIL323, rOCIL1.3kb and mOCIL cDNA sequences with the SwissProt protein database indicated that the mOCIL protein sequence included an 80 amino acid C-lectin type motif, from positions 92 to 171 in the mOCIL protein sequence. This C-lectin motif is similar to that of CD69, a membrane-bound lectin expressed by bone marrow haematopoietic cells, and thought to be involved in monocyte differentiation. C-lectin motifs are also involved in cell-cell contact and lipid binding (Sharon and Lis, 1995; Gabius 1997; Kieda, 1998).

MOCIL47 shows 100% identity to a cDNA clone encoding a C-type lectin expressed in bone marrow cells, which was deposited in the Gen Bank data base on 29 March 1999 (Accession No. AF121352). The full length mOCIL cDNA (SEQ ID NO:12) is 92% homologous to this.

The three different sequences (SEQ ID NO: 12, 15 and 17), which overall have 90% identity, may represent gene duplications, where either one or all three sequences may have similar biological outcomes. The functional data

we have to date, relating to the inhibition of osteoclast formation from haemopoietic precursor cells using antisense oligonucleotides (SEQ ID NO: 1 and 13), have been obtained mainly with mOCIL17 (SEQ ID NO: 7).

Example 3 Isolation of Human cDNA Encoding the Inhibitory Factor

[^{32}P] α-dCTP labelled rOCIL402 was used to probe a human fetal cDNA library under low stringency hybridisation conditions at 55°C in a hybridisation buffer containing 4 x SSPE, 5 x Denhardt's solution, 0.5% SDS for 24 hr, and then washed with low stringency at 2 x SSC with 0.1% SDS at 40°C for 15 min, and 1 x SSC/0.1% SDS at 40°C for 15min. Eight positive clones were obtained after tertiary screening. Clone No. 6 is a 1.3 kb cDNA segment, whose sequence was designated hOCIL (SEQ ID NO: 19). The putative protein sequence encoded by bp883-1059 was a C-type lectin moiety, which showed 73% homology to the C-type lectin sequence previously demonstrated in rOCIL323, rOCIL1.3kb and mOCIL2kb. However, regions of amino acid sequence 5' and 3' to this C-type lectin domain were different from those of the mouse and rat sequences, as shown in Figure 6.

Clone No. 8 is 960 bp long. It has a 64% identity over a length of 145 bp with rOCIL402. A search of the EST database showed that clone No. 8 has 99.5% sequence identity with a human EST clone. Accession No. AA029932, over a published length of 209 bp. This EST clone was ordered and further sequenced. The EST clone is 680 bp in length, and has 64% identity with rOCIL1.3kb over a length of 343 bp. It also has 64% identity over a length of 346 bp compared to mOCIL. RT-PCR showed that clone NO. 8 and AA029932 represent overlapping clones, which are contiguous, and combine to represent a human OCIL (hOCIL) of 1336 bp in length (SEQ ID NO: 20).

The deduced protein sequence showed 56% homology to the deduced protein sequence of rOCIL1.3kb, and 62%

homology to that of mOCIL, as shown in Figure 7. These differences are principally at the N-terminal. Although there is 80% homology between the mouse and human OCIL proteins in certain regions, this indicates that the mouse cDNA could not reliably be used to isolate a human genomic DNA encoding hOCIL.

In order to obtain the hOCIL gene, the 680 bp cDNA insert of clone AA029932 was isolated and screened by Genome Systems, Inc. against the genomic BAC Human Release II Hybridisation library, as described in Example 2. One positive clone was obtained and 7kb of this clone was sequenced. This sequence has 100% identity to a human genomic clone, which was deposited in the GenBank database on April 6, 1999 (Accession No. AC007068). These sequences, when combined, indicate that the hOCIL gene is greater than 10 kb in length (SEQ ID NO: 21). The genomic sequences corresponding to the 5' flanking region, promoter region and the first 654 bp of cDNA sequences are not represented in the sequence deposited in Accession No. AC007068.

The hOCIL gene is located in chromosome 12,12p. Chromosome 12 and chromosome 11 are considered to be evolutionarily related. There are several examples of evolutionarily related proteins whose genes are located on chromosome 12 and chromosome 11 such as PTH and PTHrP, IGF and IGF I , Harvey ras sarcoma 1 and Kisten ras sarcoma 2 etc. (Martin et al., 1991). Thus chromosome 11 and chromosome 12 share genes of similar biological characteristics with redundant function.

Example 4 Effect of Anti-Sense Oligonucleotides on Transcription

Primary mouse calvarial osteoblasts were cocultured with mouse bone marrow cells to generate mononuclear and multinucleate osteoclasts. Staining for tartrate-resistant acid phosphatase (TRAP), performed using

a commercial leukocyte acid phosphatase kit from Sigma Diagnostics (St. Louis, MO, USA) (Katsogiannis *et al*, 1998), identified these cells as osteoclasts. Under normal conditions, multinucleate functional osteoclasts are not formed unless the cocultures are stimulated with 1,25-dihydroxyvitamin D_3 and PGE_2.

Experiments were carried out to block translation of OCIL mRNA in order to determine the function of its translated product. Primary mouse calvarial osteoblasts were treated with antisense oligonucleotides. Four antisense oligonucleotide sequences were designed. Two of these antisense oligonucleotide sequences were complementary to the C-type lectin region, and designated 323 (SEQ ID NO: 22) and 402 (SEQ ID NO: 23) respectively:

323 5′-GAG TGT TGT CTG TCC ACT TCC-3′

402 5'-TTT CCA ACT CCA ATC CAG TTT-3'

The 323 antisense oligonucleotide has 19 of 21 bp complementary to mOCIL47.

```
5′-GAGTGTTGTCTGTCCACTTCC-3′ 323 antisense
   |||||||| ||||||||||
3′-GTCACAACAAACAGGTGAAGG-5′ mOCIL47 Strand +
```

and has 20 of 21 bp complementary to mOCIL and mOCIL17:

```
5′-GAGTGTTGTCTGTCCACTTCC-3′ 323 antisense
   ||||||||||||||||||||
3′-GTCACAACAGACAGGTGAAGG-5′ mOCIL17 Strand +
```

The 402 antisense oligonucleotide has 100% complementarity to mOCIL and has 20 of 21 bp complementary to mOCIL47 and mOCIL17:

```
5'-TTTCCAACTCCAATCCAGTTT-3' 402 antisense
   ||||||||||||| ||||||||
3'-AAAGGTTGAGGTCAGGTCAAA-5' mOCIL47 Strand +

5'-TTTCCAACTCCAATCCAGTTT-3' 402 antisense
   ||| ||||||||||||||||||
3'-AAAAGTTGAGGTTAGGTCAAA-5' mOCIL17 Strand +
```

The other two antisense oligonucleotide sequences, respectively designated 439 (SEQ ID NO: 24) and 474 (SEQ ID NO: 25), specifically inhibit the translation of mOCIL17.

The oligonucleotide 439 is antisense to the sense primer OCILm17 and located upstream of the open reading frame:

439 5' GAG GAG CTG AGT TTC CAC TAC-3'

The antisense oligonucleotide 474 is complementary to a region in mOCIL17 in the open reading frame but outside the C-type lectin region:

474 5'-GGT AGG GAA GCC TTT GTG AC-3'.

Under basal conditions, *ie.* in the absence of stimulation with 1,25-dihydroxyvitamin D_3 and PGE_2, there was a 3- to 5-fold increase in the number of mononucleate TRAP-positive cells in the cocultures treated with the 323 and 474 antisense oligonucleotides over the period from 3 to 7 days. With the 402 antisense oligonucleotide, a 4.5-fold increase in the formation of mononucleate TRAP positive cells was observed after 7 days treatment. Multinucleate TRAP-positive cells (323; 4.5 ± 2, 474; 4.25 ± 1.25) were also observed in cocultures treated with both 323 and 474 antisense oligonucleotides at a concentration of 5μM, whereas none were observed in the

control. These experiments were performed three times, and a representative result is shown in Figure 9a.

When the cocultures were stimulated with 1,25-dihydroxyvitamin D_3 and PGE_2, multinucleate TRAP-positive osteoclasts were formed after 7 days. Treatment with 5 μM 323 antisense oligonucleotide resulted in a seven-fold increase in the number of multinucleate osteoclasts, as shown in Figure 9b.

Treatment with 10 μM 402, 5 μM 439 and 474 antisense oligonucleotide resulted in a 2 to 3-fold increase in the formation of multinucleate osteoclasts after 7 days, as shown in Figures 9c and 9d.

These TRAP-positive cells were further characterized as osteoclasts by the presence of receptors for calcitonin, demonstrated using autoradiography and immunostaining, and by the ability of these cells to form resorption pits in bone slices.

Effects of mOCIL antisense oligonucleotides on the three phases of osteoclast formation were also investigated. Mouse bone marrow and primary osteoblastic cells were cocultured in the absence of 1,25-dihydroxyvitamin D_3 and PGE_2 for a 7 day culture period. 323 and 474 antisense oligonucleotides were added for 3 phases of culture: the first phase (0-3 days), in which there is proliferation of osteoclast progenitors, the second phase (3-5 days) and the final phase (5-7 days), in which these cells differentiate into mature osteoclasts. TRAP-positive osteoclasts were counted. In order to examine the role of OCIL on the bone resorptive activity of mature osteoclastic cells, the cells were also cultured on dentine slices under the same culture conditions as above, and resorption pits formed on dentine slices were quantitated. The results are shown in Table 1, and indicate that the OCIL acted at an early stage in osteoclast formation.

Table 1

Effects of mOCIL Antisense Oligonucleotides on the Three Phases of Osteoclast Development

TIME	TREATMENT	MONO	MNC	PITS
7 days	control	3596 ± 511.5	0	10 ± 4.4
0-3 days	323	6880 ± 674 *	7.7 ± 5.3	34 ± 15
	474	6893 ± 429.6 **	8.7 ± 1.7 *	20 ± 2.5
3-5 days	323	2840 ± 197.6	0	14 ± 6
	474	3110 ± 334	4.3 ± 2	24 ± 2.5 *
5-7 days	323	4236 ± 518.6	0	10.6 ± 0.79
	474	3363 ± 139.8	0	3.3 ± 2
0-7 days	474	5223 ± 571 *	3.3 ± 1	20.3 ± 5.9

mono mononuclear osteoclast precursors

MNC multinucleate osteoclasts

pits resorption pits formed on dentine slices

*$p < 0.05$ vs. control

**$p < 0.01$ vs. control

Example 5 Regulation of expression of OCIL mRNA

The regulation of OCIL mRNA expression was examined in the UMR106 parental osteoblast-like cell line using rOCIL402 as a probe. As shown in Figure 10, expression of the mRNA was upregulated by retinoic acid (RA), parathyroid hormone (1-34), parathyroid hormone related protein (1-34), TNF-α, interleukin 1α (IL-1 α), IL-1β, IL-11, IL-17, GM-CSF, M-CSF, TGFβ, dexamethasone, 1,25-dihydroxyvitamin D_3 and prostaglandin E_2. A time course study, illustrated in Figure 11, showed that parathyroid hormone-related protein (1-34) increased levels of OCIL mRNA as early as 1 hour, peaking at 4 hours and maintaining the high level of expression until 48 hours.

As shown in Figure 12, in primary mouse calvarial osteoblasts, OCIL mRNA was upregulated by IL-1 α, IL-1β IL-11, 1,25-dihydroxyvitamin D_3 and retinoic acid. In ST2 mouse stromal cells, OCIL mRNA was upregulated by dexamethasone, 1,25- dihydroxyvitamin D_3 and IL-11. The time course also showed that dexamethasone increased OCIL mRNA at 1 hour, peaking at 2 hours and returning a basal level at 24 hours. These results are illustrated in Figures 13A and 13B, respectively.

Example 6 Localization of OCIL mRNA

mRNA encoding OCIL was localised in fetal, newborn and adult mouse tissue by *in situ* hybridisation using the rOCIL 402 antisense probe, using a method described previously (Katsogiannis et al, 1997 and 1998). Plasmid cDNA was labelled with digoxigenin (DIG) using an RNA labelling kit (Boehringer-Mannheim, Mannheim GmbH, Germany). Hybridisation signals were detected by alkaline phosphatase staining with BCIP/NBT after incubation with an anti-digoxigenin antibody coupled to alkaline phosphatase. The mRNA is expressed in a range of tissues, as summarized in Table 2.

Table 2

Adult Rat Tissues Probed with rOCIL402

	Tissue	OCIL mRNA
Calvaria	Osteoblasts	+++
	Marrow hematopoietic cells	++
	Megakaryocytes	++
Kidney	Medulla (collecting tubules)	+++
	Outer cortex (collecting tubules only)	+
	Glomeruli (endothelial cells only weakly positive)	-ve
	Proximal/distal tubules	-ve
Lung	pneumocytes	++
	bronchial epithelium	++
Brain	Neurones in cerebral cortex, cerebellar cortex, hippocampus; choroid plexus	+++
Heart	Cardiac muscle	+++
Spleen	White pulp	+++
	Cortex	+
	Red pulp	-ve
Gut	Luminal epithelium	++
Liver	Hepatocytes	-ve

In situ hybridization was also carried out to detect OCIL mRNA localisation in adult rat tissue and human skin, using the same method, and the results are summarized in Table 3.

Table 3

Normal Murine Tissues Expressing OCIL mRNA

Tissues	Fetal (day 15)	Newborn (day 1)	Adult (5-8 weeks)
Extraskeletal tissues			
Brain	+++	+++	+++
Lung	++	+++	-
Heart	+++	+++	++
Kidney (collecting tubules)	++	+++	-
Small Intestine	+	+	-
Liver	+/(mk=+++)*	+	-
Skeletal muscle	+++	+++	++
Skin	+++	+++	++
Spleen	nd	nd	++
Skeletal tissues/cells			
Long bone			
chondrocytes	+++	++	+++
osteoblasts	na	+++	+++
osteoclasts	nd	++ or -	++ or -
perichondrium/ periosteum	++	+++	++
marrow/megakaryocytes	na	++	++
Calvarial bone			
osteoblasts	+++	+++	++
osteoclasts	++	++	++
periosteum	nd	++	++

(+) denotes weak signal
(++) denotes moderate signal
(+++) denotes strong signal
(-) denotes absence of signal
(na) not applicable

(nd) not determined.
*(mk) megakaryoblast of fetal liver.

OCIL mRNA localization in human skin probed with OCIL402:

Epidermis (all layers) +++
**Basal layer slightly weaker signal

In Northern blot analyses of adult mouse tissues using mOCIL2 as the probe, OCIL mRNA was shown to be expressed in heart, skin, lung, liver, kidney, gut and brain. In adult rat, OCIL mRNA was found to be expressed in brain, bone, lung, liver, gut, kidney, mouse, skin and heart. These results are illustrated in Figure 14.

Example 7 Antibody Directed Against OCIL

The following peptide fragment of the deduced amino acid sequence derived from the cDNA sequence of OCIL was synthesized, and was used to immunize rabbits, using standard protocols.

H-Cys-Met-Arg-Thr-Glu-Ala-Gln-Leu-Ala-Arg-Phe-Asp-Asn-Gln-Asp-Glu-Leu-Asn-Phe-OH (SEQ ID NO: 26)

The antibody raised may be used to detect the OCIL protein, using standard immunohistochemical methods, or to neutralize OCIL activity in murine co-cultures to stimulate osteoclast formation.

Example 8 Immunohistochemistry

A rabbit polyclonal antibody prepared as described in Example 7 was used for immunohistochemistry. A kit for the standard peroxidase-labelled streptavidin-biotin detection method (DAKO, Boenisch, 1989) was used according to the manufacturer's instructions with minor modifications. The dilution of the antiserum used was optimised in preliminary experiments. Incubation of tissue

sections with a 1:100 dilution of the primary antiserum was carried out overnight at 4°C in a humidified chamber. Peroxidase activity was detected with 3'-3'-diamino-benzidine tetrahydrochloride (Sigma) and 0.15% H_2O_2. Slides were counterstained with haematoxylin, dehydrated and mounted on a coverslip. The results are summarized in Table 4.

Table 4

Normal Murine Tissues Expressing mOCIL protein

Tissues	Fetal (day 15)	Newborn (day 1)	Adult (5-8 weeks)
Extraskeletal tissues			
Brain	nd	+	+++
Lung	nd	++	+
Heart	nd	++	++
Kidney (collecting tubules)	nd	++	++
Small Intestine	nd	nd	nd
Liver	nd	nd	nd
Skeletal muscle	nd	++	++
Skin	nd	++	++
Spleen	nd	nd	nd
Skeletal tissues/cells			
Long bone chondrocytes	nd	++	++
osteoblasts	na	+++	+++
osteoclasts	nd	nd	nd
perichondrium/periosteum	nd	++	++
marrow/megakaryocytes	na	++	++

(+) denotes weak signal;

(++) denotes moderate signal;

(+++) denotes strong signal;

(−) denotes absence of signal;
(na) not applicable;
(nd) not determined.

Example 9 Recombinant OCIL protein

OCIL proteins were prepared by recombinant DNA technology to allow more extensive laboratory studies of their actions on osteoclast formation as well as osteoblast function. Soluble mouse and rat OCIL cDNA tagged at the N-terminus with the FLAG epitope were constructed in the pEF-BOS Mammalian expression vector (Mizushima & Nagata 1990), which had been modified to contain an in-frame IL-3 signal sequence and FLAG peptide coding sequence (gift of Dr. D Hilton).

In order to obtain a RT-PCR product encoding the mOCIL extracellular domain (amino acids 77-218) to clone into the MluI site of the vector, as shown in Figure 6, the RT-PCR was carried out using total RNA isolated from primary mouse calvarial osteoblasts, which support osteoclast differentiation in coculture. A sense primer, OCILm33, comprising OCILm32 representing nucleotides 243-267 of mOCIL and containing a MluI site, designated primer OCILm33 (SEQ ID NO: 27):

OCILm33 5′-GCC ACG CGT TTG TCA GCA ACA AAG ACA GAA CAG-3′

and an antisense primer comprising OCILm12 (SEQ ID NO: 4), representing nucleotides 743-720 of mOCIL and containing a MluI site, designated primer OCILm46 (SEQ ID NO: 28),

OCILm46 5′-GCC ACG CGT GGG ACC ATA GGG GAA AAA GTA G-3′

were used as primers in the PCR. PCR was run at 94°C for 5 min, then 35 cycles of 94°C for 30 s, 60°C for 30s, and 72°C for 1 min, followed by a final extension step of 72°C for 10 min. A 501 bp fragment was obtained and further cloned into the expression vector pEF-BOS. The open

reading frame and FLAG fusion was confirmed by sequencing (bp 1-132), and the 501 bp fragment sequence (SEQ ID NO: 29) was confirmed to be identical to mOCIL17.

To obtain a PCR product encoding the rOCIL1.3kb extracellular domain (amino acids 40-179), a sense primer to represent nucleotides 126-146 of rOCIL1.3kb with the MluI site, designated primer OCILr22 (SEQ ID NO: 30),

OCILr22 5'- GCC ACG CGT TCA GTA AAA AAG ACA GCC AAG-3'

and an antisense primer representing nucleotides 544-526 of rOCIL1.3kb with the MluI site, and designated primer OCILr23 (SEQ ID NO: 31),

OCILr23 5'-GCC CAG CGT AAC TAC AGG CAC TGT GAG G-3'

were used as primers in a PCR. PCR was carried using rOCIL1.3 kb plasmid as template and run at 94°C for 5 min, then 35 cycles of 94°C for 30 s, 60°C for 30s, and 72°C for 1 min, followed by a final extension step of 72°C for 10 min. A 421 bp fragment was obtained and cloned into the expression vector pEF-BOS. The open reading frame and FLAG fusion were confirmed by sequencing.

HEK 293 cells were transfected with both mouse and rat expression constructs using Lipofectamine (Life Technologies, Inc). Supernatant was harvested after 72 hours. The recombinant protein was purified by incubation with the anti-FLAG M2 affinity gel (Kodak), and eluted with the FLAG peptide (Kodak) as outlined in the manufacturer's protocol. The purified protein was used to study its effects on osteoclast formation in murine cocultures.

An experiment was also carried out to determine the action of rOCIL protein in osteoclast formation. Primary mouse calvarial osteoblasts were cocultured with mouse spleen cells obtained from 15 week-old mice and stimulated with 1,25-dihydroxyvitamin D_3 and PGE_2 in the

presence of rOCIL protein for 10 days. A negative control was carried out as with protein carrier buffer alone. As shown in Figure 15, rOCIL protein significantly reduced the number of osteoclasts formed when compared to the presence of carrier buffer alone.

DISCUSSION

We conclude from our experiments that in osteoblasts OCIL is expressed on the cell surface as a type II membrane peptide. Contact with haematopoietic precursor cells prevents further differentiation into mononucleate osteoclast precursors, and ultimately into functional multinucleate osteoclasts. Without wishing to limit the scope of the invention by any proposed mechanism, we consider that upregulation of OCIL mRNA expression by the same osteotropic factors that increase expression of ODF is consistent with the hypothesis that regulation of bone resorption by osteoclasts is tightly regulated. According to this hypothesis, stimulation of multinucleate osteoclast formation through ODF would simultaneously prevent the generation of new osteoclasts through the action of OCIL. If this system is operative under normal physiological conditions, then bone resorption becomes a self-limiting process.

In vivo, OCIL has the potential to be used as a therapeutic agent in the treatment of conditions which are characterised by excessive bone resorption, such as osteoporosis, primary hyperparathyroidism, Paget's disease, rheumatoid arthritis, renal osteodystrophy, and humoral hypercalcaemia of malignancy, as well as metastatic bone disease. Modulation of the expression or function of the factor may also be useful in the treatment of disorders involving extra-skeletal calcification.

It will be apparent to the person skilled in the art that while the invention has been described in some detail for the purposes of clarity and understanding,

various modifications and alterations to the embodiments and methods described herein may be made without departing from the scope of the inventive concept disclosed in this specification.

5

References cited herein are listed on the following pages, and are incorporated herein by this reference.

REFERENCES

Anderson, D.M., Maraskovsky, E., Billingsley, W.L., Dougall, W.C., Tometsko, M.E., Roux, E.R., Teepe, M.C., DuBose, R.F., Cosman, D. and Gaiibert, L.
Nature, 1997 390 175-179.

Boenisch, T.
In: Naish, S.J., Ed Handbook: Immunochemical Staining Methods. DAKOPATTS (DAKO Corporation): CA; 1989 16-17.

Chomczynski, P. and Sacchi, N.
Anal Biochem, 1987 162 156-159.

Ganius, H.J.
Eur. U. Biochem., 1997 243 543-576.

Kartsogiannis, V., Moseley, J., McKelvie, B., Chou, S.T., Hards, D.K., Ng, K.W., Martin, T.J. and Zhou, H.
Bone, 1998 22 189-194.

Kartsogiannis, V., Udagawa, N., Ng, K.W., Martin, T.J., Moseley, J. and Zhou, H.
Bone, 1997 21 385-392.

Kieda, C.
Adv. Exp. Med. Biol., 1998 435 75-82.

Martin, T.J., Moseley, J.M. and Gillespie, M.T.
Critical Reviews in Biochemistry and Molecular Biology, 1991 26 377-395.

Martin, T.J. and Udagawa, N.
Trends in Endocrinology and Metabolism, 1998 9 6-12.

Mizhashi, N., and Nagata, S.
Nucleic Acids Res., 1990 18 5322

Ng, K.W., Gummer, P.R., Michelangeli, V.P., Bateman, W., Mascara, T., Cole, W.G. and Martin, T.J.
J Bone Miner Res, 1988 3 53-61.

Sambrook, J., Fritsch, E.F. and Maniatis, T.
Molecular Cloning - A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, 1989.

Sharon, N. and Lis, H.
Essays Biochem., 1995 30 59-75.

Simonet, W.S., Lacey, D.L., Dunstan, C.R., Kelly, M., Chang, M.S., Luthy, R., Nguyen, H.Q., Wooden, S., Bennett, L., Boone, T., Shimamoto, G., DeRose, M., Elliott, R., Colombero, A., Tan, H.L., Trail, G., Sullivan, J., Davy, E., Bucay, N., Renshaw-Geee, L., Hughes, T.M., Hill, D., Pattison, W., Campbell, P., Boyle, W.J. *et al*
Cell, 1997 89 309-319.

Suda, T., Udagawa, N., Nakamura, I., Miyaura, C. and Takahishi, N.
Bone, 1995 17 87S-91S.

Takahashi, N., Akatsu, T., Udagawa, N., Sasaki, T., Yamaguchi, A., Moseley, J.M., Martin, T.J. and Suda, T.
Endocrinology, 1988 123 2600-2602.

Tsuda, E., Goto, M., Mochizuki, S., Yano, K., Kobayashi, F., Morinaga, T. and Higashio, K.
Biochem. Biophys. Res. Commun., 1997 234 137-142.

Wong, B.R., Rho, J., Arron, J., Robinson, E., Orlinick, J., Chao, M., Kalachikov, S., Cayani, E. and Barlett, F.S. 3rd, Frankel WN, Lee SY and Choi Y.
J. Biol. Chem., 1997 272 25190-25194.

Yasuda, H., Shima, N., Nakagawa, N., Yamaguchi, K., Kinosaki, M., Mochizuki, S., Tomoyasu, A., Yano, K., Goto, M., Murakami, A., Tsuda, E., Morinaga, T., Higashio, K., Udagawa, N., Takahashi, N. and Suda, T.
Proc. Natl. Acad. Sci. USA, [1998] 95 3597-3602.

ST VINCENT'S INSTITUTE OF MEDICAL RESEARCH

SEQUENCE ID LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: ST VINCENT'S INSTITUTE OF MEDICAL RESEARCH

(B) STREET: 41 VICTORIA PARADE

(C) CITY: FITZROY

(D) STATE: VICTORIA

(E) COUNTRY: AUSTRALIA

(F) POSTAL CODE (ZIP): 3065

(ii) TITLE OF INVENTION: INHIBITOR OF OSTEOCLAST PRECURSOR FORMATION

(iii) NUMBER OF SEQUENCES: 31

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATGCTGGGCA CGTACACACA A 21

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 321 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

CGCTCTAGCC CGGCCACGCG TCGACTAGTA CAGCTCCAAA TCTGTGCCCC TCAGTTCCTC 60

CCTCCTGTTA TCTCTAGAGG AAGCTGTGGA GAGATTCCAG GATCATCTGA AACAGAGACA 120

CATGCATTCT CGGCTTTTTG TGTTTTATTA CAGAATTTCT TAAGCAGATA CAAAGGGAGT 180

TTTGATTACT GGATCGGCCT GCACAGAGAG TCCTCAGAGC ACCCTTGGAA GTGGACAGAC 240

AACACTCAGT ATAACTACTC GTATGTTTCA CAATGTTTTT TCTTCTACTG TGTTCATGTC 300

TTGTTGAGGT CTTGTGTGTA C 321

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TGAGTGTTGT CTGTCCACTT CCAA 24

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 402 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

ACAGTAAAAT GCTCCAAGGA AAGCTTCCCA GAAACATCCC CCTGGAGTAT CCTGCTGGGC 60

CTTACTGCTG CTACGTAGTG ATCATTGTCC TCAGTGTTAG CTGTAGTTCT CTTTCTGTTG 120

CTTTGTCAGT AAAAAAGACA GCCAAGATCT CAACCATAAA TACTTATGCT GCTTGCCCGA 180

GAAACTGGAT TGGAGTTGGA AATAAATGTT TTTATTTTAA TGAAATACCA AGTAACTGGA 240

CATTGAGCCA GACCCTCTGT AAGGAACAAG GGGCCGAGCT AGCACGATTT GACACCGAGG 300

AGGAGCTGAA TTTCCTAAGG AGATACAAAG GGAGTTCAGG TTACTGGTCC GGTCTGCACA 360

GAGAGTCATC AGCGCACCCT TGGAAGTGGA CAGACAACAC TC 402

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GAAACATCCC CCTGGAGTAT CC 22

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CCAAGTAACT GGACATTGAG CCAGA 25

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1302 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

ACAGTAAAAT GCTCCAAGGA AAGCTTCCCA GAAACATCCC CCTGGAGTAT CCTGCTGGGC 60

CTTACTGCTG CTACGTAGTG ATCATTGTCC TCAGTGTTAG CTGTAGTTCT CTTTCTGTTG 120

CTTTGTCAGT AAAAAAGACA GCCAAGATCT CAACCATAAA TACTTATGCT GCTTGCCCGA 180

GAAACTGGAT TGGAGTTGGA AATAAATGTT TTTATTTTAA TGAAATACCA AGTAACTGGA 240

CATTGAGCCA GACCCTCTGT AAGGAACAAG GGGCCGAGCT AGCACGATTT GACACCGAGG 300

AGGAGCTGAA TTTCCTAAGG AGATACAAAG GGAGTTCAGG TTACTGGTTC GGTCTGCACA 360

GAGAGTCATC AGCGCACCCT TGGAAGTGGA CAGACAACAC TGAGTATAAC AACTCGGTTT 420

CCATCGGAGG AGATGAAAAA CATGGCTTCC TGAGTGACAA TGGGTTCAGC AGTGGCAGGG 480

GTTATATAGT GAGGAAGTCG ATTTGTAGGA AGCCCAACAG CTACACCTCA CAGTGCCTGT 540

AGTTTTGTGT CCTTGGTTGA GACTTTGTCC TAACAGTCAT GAGGAACACA GAACATGGTA 600

TCTACAGTGC CTGAATCATG AACAATCTGC TAAAATCATC TTCAATTCAT AATGTGTGGT 660

GACATCTAAG ATAACAACTG AGGCATATTT TGCTTGGGAG ATCATGAATT GTTCTATATT 720

AAATAGGTAT TCAGGTATGA GCTGGTTCTC ACATCTTAAA CATAAACTGA ATCATGTCAG 780

TATTAGTTAT CTCTACTTTC TTTTTTCTCT CATTTAAATT ATATTATTTA TTTATATTCC 840

AAATACCGTC CCCTCCTTGT TCCCCCTTCT AGAGTTGTTC ACTCCATACC CCTTCATCTT 900

TACTTCTGAA GAGATGTTCC CCCACCCCAC TCTGAGTATT TCCCTTCTCT TGGACTTTAG 960

GACTGTACAG GATTAGGTGC ATCCTCTCAT AGTGAGGCCA ACTGTAGGGA GCTGCGACAT 1020

GCCGTGCCTC AAAATGGTGC TGGTTTCCGC CTTCCACCCT CCCAACAGTG AGCGCTCCTT 1080

GTAGTAAACA AGTCCTTATT TGACTATGCC TGCCTGGCCT GCTAGGTTCA GCATAGTGAC 1140

AGCCTGTCTG CATGACCCAT GTGGCACGTT GGGGTTGGTT GGTGTTGGAT ACATAAGCTG 1200

ATGTAGGGCA TTCCCCTGGG GTAGTAGATG ATTGTATCAA GGTTCCTGAA TAAACTGCTT 1260

GAAGAAAAAA AAAAAAAAAA AAGTACTAGT CGACGCGTGG CC 1302

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 738 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

AGTAAAATGC TCCAAGGAAA GCTTCCCAGA AACATCCCCC TGGAGTATCC TGCTGGGCCT 60

TACTGCTGCT ACGTAGTGAT CATTGTCCTC AGTGTTAGCT GTAGTTCTCT TTCTGTTGCT 120

TTGTCAGTAA AAAAGACAGC CAAGATCTCA ACCATAAATA CTTATGCTGC TTGCCCGAGA 180

AACTGGATTG GAGTTGGAAA TAAATGTTTT TATTTTAATG AAATACCAAG TAACTGGACA 240

TTGAGCCAGA CCCTCTGTAA GGAACAAGGG GCCGAGCTAG CACGATTTGA CACCGAGGAG 300

GAGCTGAATT TCCTAAGGAG ATACAAAGGG AGTTCAGGTT ACTGGTCCGG TCTGCACAGA 360

GAGTCATCAG CGCACCCTTG GAAGTGGACA GACAACACTC AGTATAACTA CTCACAGAGC 420

CTCAGATGGG GAGCCGGGAC TCTGAAATCC CAGAAAGCCA CTGCAGAACT GCAAGCCTGA 480

GATTTTGATG TCCACTATTT GCATGGCTGC ACCTGTTCAG GAAAGCAGAG ATTTTAAGGA 540

CATTCGGAAC CTCCTTTAAA GTTTTGTCAT CACAGAGCAC CCAAAACAGT CCTCGAATCA 600

CAGGCCCAGT CCCATCCACC GTTAAAGCAC CTTTGAGCAA TTTAATAAGA AGTGCGTGTT 660

CCCATGTGTA AAATGAATAA AAACAGAATT GGAAAAAAAA AAAAAAAAAA AAAAAAAAAA 720

AAAAAAAAAA AAAAAAAA 738

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 620 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

AGTAAAATGC TCCAAGGAAA GCTTCCCAGA AACATCCCCC TGGAGTATCC TGCTGGGCCT 60

TACTGCTGCT ACGTAGTGAT CATTGTCCTC AGTGTTAGCT GTAGTTCTCT TTCTGTTGCT 120

TTGTCAGTAA AAAAGACAGC CAAGATCTCA ACCATAAATA CTTATGCTGC TTGCCCGAGA 180

AACTGGATTG GAGTTGGAAA TAAATGTTTT TATTTTAATG AAATACCAAG TAACTGGACA 240

TTGAGCCAGA CCCTCTGTAA GGAACAAGGG GCCGAGCTAG CACGATTTGA CACCGAGGAG 300

GAGCTGAATT TCCTAAGGAG ATACAAAGGG AGTTCAGGTT ACTGGTCCGG TCTGCACAGA 360

GAGTCATCAG CGCACCCTTG GAAGTGGACA GACAACACTC AGTATAACTA CTCGCTTTCC 420

ATCCGGGGAG TGGAAAGATA TGCCTACCTG AACGACATCG GGATCAGCAG TGCCAGGGTC 480

TATGCAGACA AAAGATGGAG CTGTAGCAAA CTTAACAGCT ATAGCCTCCA ATGCAAAACT 540

CCTTTTTCTC CTATGTAGCT TTTGATCAAG AGAGATGCTT TTTAGTCTGC TAAAAAAAAA 600

AAAAAAAAAA AAAAAAAAAA 620

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1907 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CCGAATGTTT CCTGCAACAC AAAGATGACA ACCCCAGCCT GCCACCATTT GAAAGGCCAG 60

AGGCTGAGGC CATGTGCACC TTCCATTTCA TTTCTGATGT TAAGAAATAT TCTCTATCTG 120

GTTTGATAGC ACTTTGGGAC CATAGGGGAA AGAGTAGCAC CCACAGATAA CAGGCTAAAA 180

AGCGTCTCTT GGTAAATGCT AGGAAGGAAA AAAAGGAGTT TGGCAGTGGA GGCTATAGCT 240

GTTGAGCTTG CTACAGATCC ACATCCGAAG TGAATAGATC CTGGTACTGC TGATCCCGTT 300

GTTGTTCAGG TAGGCAAATC TTTCCTCTCC CCGGATGGGA ATCGTGTTGT TATACTCAGT 360

GTTGTCTGTC CACTTCCAAG GGTGCTCTTT GGCCTCTCAG CTTTCAAGTT TCAATCCTGT 420

AGTGGAAACT CAGCTCCTCA GCTCTGAGAT GTGTGTCACA AAGGCTTCCC TACCTATGCT 480

TAGTCCCACA GGCAGCCCGC AGGTAGAAGT GGGTAAAATT CTCCAAGGAA AAAGGCACGG 540

AACCATCTCC CCTGAGTCTT GTGCTAAGCT TTACTGCTAC TATGGAGTGA TCATGGTCCT 600

CACTGTAGCT GTAATTGCTC TTTCTGTTGC TTTGTCAGCA ACAAAGACAG AACAGATCCC 660

AGTCAACAAG ACCTATGCTG CTTGCCCGCA AAACTGGATT GGAGTTGAAA ATAAATGTTT 720

TTATTTTTCT GAATACCCAA GTAACTGGAC ATTCGCCCAG GCCTTCTGCA TGCGCACAGA 780

GGCCCAACTA GCTCGGTTTG ACAACCAGGA TGAGCTGAAT TTCCTAATGA GATACAAGGC 840

GAATTTTGAT TCCTGGATTG GCCTGCACAG AGAGTCGTCA GAGCACCCTT GGAAGTGGAC 900

AGACAACACT GAGTATAACA ACACGATTCC ATCCGGGGAG AGGAAAGATT TGCCTACCTG 960

AACAACAACG GGATCAGGGA ATTCCGGGAC ACCCGTCAGC ATTCCTGGAG AAAATTCGGC 1020

ATTCATGAGA AAACTGTCTT TCTACTCCAG TGCTCTCAGT GACCAATGGC TACTGAGTGC 1080

TGCTTCATCT GAACTGATCT GAATTGAGGC AAATGTAGGG TTGGCTTCCT GCAGGAAGAC 1140

TGTTCAAAGC CAAGCTCTTT CCCTTCTAGG TGCCTGGGTC TAGTGCACAT TAGTCTTGTT 1200

GGCAGCGTGT CTCCTCAGTC TGGCTATTGT GATCTTTCCC ATAGAAAGAG TCAGGAACGA 1260

GGGGAAGGGA AAGATAGAGG CCTAAGGTGA AATTTTAAAA AACTCAATCT GTTGGTTTGA 1320

TTTGTGGTTT CATGTTTGGG TGCAATTGTT CTTGAGACAA AAGTAGAACT TTGAAATACT 1380

TTATTTAAAG AAACGAGTGC TCTGGCATTA TTAAATAAAC CTAATGTAAG TCTATGAAGA 1440

GTTTCACTTA AATACATTTA TATAAAGAGC CAATGTTAAA AGTGTTATGG ATAATAATTC 1500

TTCAAGGTGG TGGTTGTATT GGAACAAGTG TTCTTTCTGT CAGCTAGATT CCTGGTATAA 1560

AATAATTTGA CTGCAGGGAA GTTGACAGAA AGCATTACTT CTGTATGCTA CAACCCTTTA 1620

AAATTGTGCT CTGCCTCCAC CCATGTGGTG GTTTGAATGA AAATGTGGCC ATAGTCTCAT 1680

ATTTGGATGT TTAATCACTA GGGAATGGAC CTGTTTGATA GGATTAGAAG GATTGGAGGC 1740

GAGGCCTATT GGAGGAAGTG CCATACTGTG GATGGCCTTT GCCTAGTCTG TCAACCCCAG 1800

AGTTTTCATG CCTGAGTGCT CCCTGCTGGA TAATGGAGTA ACCCTCTGAA ACTGTAAGCA 1860

AGCTCCTGAT TAAATGCTTT CATTTCTAAA AAAAAAAAAA AAAAAAG 1907

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9349 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

TGCATTACAC ACACACACAC ACACACACAC ACACAAGGCC GGGCAGTGGT GGTGCACACC 60

ATTAATCCCA GCACTGGGGA GGCAGAGACA GGCAGATTTC TGAGTTCAAG GCCAGCCTGG 120

TCTACAGAGT GAGTTCCAGG ATATCCAGGG CTACACAGAG AAACCCTGTT TCAAAAAAGT 180

TACTTTTTGT ACCTTGAAAT CTAAAATATG TCTCAACTCT GTTTGTTTCT TTTACAGTAT 240

AACATGCTCC CCCCCCCCCC CCGCCGCCGC AGTTTTTCAG TTCCAGATCT
AGGTAGGCAC 300

CCAATCTCTG GCAGCTTATC AAGTCAGCTG ATGTAAAAAT AATCCCACAA
CTCACAAAAT 360

ATAGAGGGAA GACAGCGGGG AAAAAGGGGC GGGCTCATTG CTTCAGCAAG
AAGATAGTGG 420

TGCATAGCCT CCCATGCCAG ATTGCTTGGA GACAGGAGAA AAACTGTACG
TATTTAATGA 480

AATGCTAACT AAACTAAAGT GGGGGAGGCT TCCTCAGGGG AGCTGGATCT
TGCTCCTGTT 540

AGCCTGCCAT AGTGGGTCTA TATAGACCAG CTGAGGCTGG GGTGGGGTGG
ATGGTGGGAG 600

CTCTGCTGTG GTCGGAAAGT ACCGATGCCA CTCTGGCTTT CTGGTATGGC
CAATGTTACT 660

TAAATACGTT TGGGAGGAGT GCAACCTTTT GAGTTTGTAA ATAAAAGCAG
GTGCCCAGAT 720

TCCTGGAGGA TTGACTGGAG GACCTTGGGG GTGCTCTGGC ACACCCTGCC
ACCCAGCCCA 780

TACCTTAAGT GCCCCTCCTA CACACCTACC TACAACTTTC TTTTCAGGCT
CCCACAGTAC 840

TCCCCCTTTC CCAAACCTCC AAGCTTTTGG AATTTCTCTC TCTTCCCAAG
GACACGGTAT 900

CAGGTAATAC TCTTTCTGGC CTTAAATGAC TCTTGTGGCA CCAGGGAAGG
ATCAGTTTTT 960

TTCCAGTAGG GTGGGGGTGG GAGATTTATC CCATCTACAA ATCCATCTAC AGTTTTAGTT 1020

CACTGGGTGC TGGGAATGAA CCAAGTCCTC TCTCTGCAAG AGCAGCAAGC TCCCTTCCCT 1080

GTTGAGCCAT GACTTTACCC CCACTTTAAT ACTTTTGTTT AGGAATAAAA TATCAATTTT 1140

CTTGAAAAGC AGAGTTCACA ATTGTTGTTA GATCAATGGC CTAGTGGCAG CCTGAGGATA 1200

CCAGGCAAGC TCCTTCAGAG TGGACAGCCT AGCTGCTAAG ATGATTGGAA ATACTGTTCT 1260

GGGAGGTGGG GGACAGGTCG AGGAAGAGGG AGACCTAACC ATGCCTCCCT TCAACCCTAG 1320

GGCCCTACTC CATGCCATCC TGTGCACACC TAAAGTACCC TCCTCCACGG CTATCCTGGT 1380

CCCTTAAACA GACCCTTAAT CAGAGTGTAG AACAGGGTCT TCTTGAGGCA GAGTAGCAGG 1440

TATGATTGGC CTGCTGCCTT TGACTGTGAG CTATAGCCAG GTTCCACCAA GTTCCATACT 1500

CCTCACAGTA AGCCATAAGC GCCTGTTTGT GTTGGGAAAA CTTAGAAAAA GTAAAGATTT 1560

TCTTTTGTTC TTGAGACTTT TCTATGGGTT AAAAATGGCG AGCCAGGGTC CTACCAGCAG 1620

TGGGTCAGGG GGACATGAAA GGGACCTGAA TTTGGTGAAA AGGTACTGGA TGTGGTGTCT 1680

CACAAGTGGG TCTTTTCTAG AAGCCCACCG CAGCTTTCTT TTTATGGTTT
TAACCTCCCT 1740

CTGGCTTAGT CCCCAAAGAG GAAAGGGCCG TGGAGGAAAT ACTAGTTGTG
TTCTTGAAAA 1800

AAAGCCATGT GCTTTTTGAG GGCAAGTAGT TTTAAAGGCA ACGCTTGAAG
GGGCCCTTAA 1860

AGAAAATCCG GGTAATGGCA GGAAGGTATT TTTATAAATT TTTGGTAAAG
GAAAATAACC 1920

TATATGGCCC CAACTGGGTT GAGTCCCAGC GTGGTCACCT TGGCTGGGGA
CAAGGGTGCT 1980

TCAGCTCCCA TGGATGTGGG AACTGGGCTT TGGGTGCCAG GGCTGGGGTG
GTGGAAGGTT 2040

TTGGGGTATG GCCCAAGTCG GAGGCTCCCA ACCAAACAGG CATCCAGCAT
CTTTCCAGTG 2100

GATGAGTGTG GAAAACCTGT GATACATACT CCCATATATA CTGGAATACT
ATGTACTAGT 2160

AAGATAGGAT GTCTTTTGTG ACAACATGGC TGGACCTGGG GTGACATGCT
GAGACAAATT 2220

AGTCAGGCAC TGAAAGGCCA ACATTGTTCA TCAGTTGTAG AGGGGTTTGT
TAGCTAAAAG 2280

CAGACAGGAG TTTACACTCT TTTCTTCGAT TTGGAAAGAT TTTTGAAATC
ACAGTGCAGA 2340

ACCTGAAATC ACAATGAAAC CAAACCACTC CTTTACAATC TGAAGGGGTT
TAGAAATCTC 2400

CCAAGACTTC CTTTCTATAG GGAGTGTGAG GAGGGCTGAG GAGGGCTCCC AGCAGCACAT 2460

GGCTGAGAGG TGCTGGGGCT GGAAATGAGC ACAGGCGAAT TTATTATGCT ATCATTTTAT 2520

ATTCTGTAGA ACTAGAAAGA ATTAAGGCTG GGAGTTCTGT GTGGATCCAA AATGCAAAAG 2580

CTCAGTGCTT AAAGCCTTCT TTCTAATCCT AAGGCTCCTT TCCCTCCTTG TTAATGTAAT 2640

AGAAGCTTTC TGGTATTTTA GGTGTGCGAA AATGCACAAA ATGCAAGGAT TAAAGTCAGT 2700

GAAAACTCTG TAAAAACTAT AATTAGCACT CAATAAAATT AATTCATTTG GTATACATTT 2760

CTGTGAATTT TGAAAACATA TAATCAGGTG TTCTTCATTA AGATACATAG GGGCTGGAGA 2820

CTTGGCTCAA CCACTGAGAG CATTTATTGC TCTTGCTGAG GACTGAGGTT TCACTCCCAG 2880

CACACATATG GTGGCTCAAC ACCCACCCCT AATTCCAATT CCAGGGATCC AATATATTTT 2940

CTAAATTCCT CTAACAGTAA TCATGCATGT AGTACACTAC ATACATACAT ACATACATTA 3000

CATTCACACA TTCTTACATT TAGCTGACAA AGCACTCTTA AATGTAAAAT AAATAAGACT 3060

AAAACAGTCA TTTTAAAAAT ATATACAGAC CCCCTACCCT ACCTGTTTCC CCGTTGTCTG 3120

CTGCAGACAC TCTCACCACT CCTCCGCCAC AGCCATGAGT AGTCACCTTT CCAGATGACT 3180

TAAAATGGGT CCATGAAGCA GAGAAGTCCC ACAAGAGTTC TTTCAGCTTG TCACAGCAAT 3240

GCCTTCTGCT CATCACTCAC AGTGCAGTGC CAATCAGTAG TGTGTCAGAA ACATGCACTG 3300

CTGGTGAGAT GCTGAGGGAT CATACCCATA GCATCGCCTA CACAGAATCA TGCTCTGAGT 3360

TCAGAAATTT TTAAGAATCT CACCAGCAAA TACTATGCAA AGAGGTTGTG AAAAGCTGTC 3420

AGGAAACTTC TAGAGAAGTG ATAGGAGGAA GTGAATAGTG GCAGTTGGGG GTCTCTTCAC 3480

AAAGGAAACT GGGACTTCCT GTAGCTCTCT GACCTTTGCA TGAGCTTACT TTCGGGTTAG 3540

TTTAGGGACA CTTTGGGGAA GAAGCCCTTG GGACATTTGG CCTGTTAAAG TGGCATGAGA 3600

TAAGGCAAGC ACAGGCATGT GTTCCAAGTT GTTTCTTGTG TTGAGAGGTT TACCTTGTCA 3660

TCAGCTTGGG GATATTTTAA TGGTCACAAA TGTGTCATTT TCACAGGGAT CCTTAAAGCT 3720

GCTGCAAATA TTCACATAAA GATGTCTTGC AACTTGAATT CCTTTCCAGC ATGGGAATAT 3780

GTGGGTAGGA TGGGAGCATA TCACATTTTA CACTTACAAA CAGCTTTGTA GAAGCTGTAA 3840

AATTTAGCCT TAAGAAGTTG TTAGTTCTAC CTCAACATGG ACATCCACAT
CAATGTATAA 3900

CCATCCTGTT ATGCAGACAG TGATTTTGCT CTTAAATCGA AGATGATTTT
GCCCAAGACA 3960

AGTTCACAAA CATTCCCTTA CTTTTCTAAA AATCAAATGA CTTTATGATA
TTAAGTTTTG 4020

TGCTTGGGAT CTCTATGTCT ACAAATGGAC TGTAGAAATT TATGCCTATT
TATTTATTTA 4080

TTTATTTTTT TAGGAGACAG ATGGAAGGGT GTTTCAGTAG CACACACGGG
GAGTGATGCC 4140

ATCTTTTCCC TGTCTAAAGA CTGGTTCATC TCTGGTTAAG TGGTCTCTTG
ACCACCCACA 4200

TGTGTTTGAC CTCATTGTGG AGTCCTGTTT TCTGCTGTGT TGTTTCAGTG
TCTACTCTGA 4260

TGCTAGCACC AGGCTTTCAT TCCTGTTCCC ATGAGAACAT GAGACATCAG
GTCAGGCCTT 4320

GTGATCTTTC TGATTTTGAC CTCTTCATTC TCAGAATAAT TTTTGACTAT
TAATTTTTGA 4380

CTCTTTTTAA TTTTCATATT ACTTCTCATA CAACTTGGTG CTATGATTTT
TTTTTTTCTG 4440

AAGGCAACTA CATCTCTGGA ATATGTTACA CATATATGTG TGTTCAGAAA
TTGTTGACAT 4500

GAGCAATATG GAGTGTTCTA GTTCATTATG TGTTTAGCGC TGCCTTGATT
TCTTTCTTTA 4560

TTTCAAAAAT GTAACATGTA CGAGTGCTTC GCCTGCATAG ATATCTATGG
ATTGTGTGCA 4620

TTCCTAGTGC CCACCAAGGT TAGGAAAGTT CTCAGACTTC CTGAATATGG
ATTTACAGAG 4680

GCTTGTGAGC TCTTATATTG GTTCTAGATT TAACCCAGGT CCTCTGAAAG
ATCAACAAAT 4740

GTTCGAAACC ACTGAGCCAA CTTTCCTTTT TATTTCTTTA TATATTTACA
TGTAGTTTTT 4800

GTTATGCTGA CTATGAACAT CCTATTTTTA GATTTGTAAT TTTATGTTTT
TTTGGCATTA 4860

CTGTCTAAGA TATTAATTTC CACTTGACAG GACAATACAA CTGATTTTAC
CATGCTGTCA 4920

CCCTGTCCTG CAGTTTTATA AACTCATTTT TTTTTTTTTT TTTACATCCA
ACGGTCTTTT 4980

GTAGATTTGT TTGGATTTTT ATGTAGATAG AACCCTGTCA TCTGTGAAAA
GAAACACTTT 5040

CACCTGTTTC TGAGAAGCAT AAACCTCCTT TCGTTCTTTA AGAGACTCAT
TTTATGTATG 5100

TGGCTGTTTG CCTGCAGGTT TGTACATAAC ATGAATGCCC AGTGTGTACA
GAAGCCAGAA 5160

GAGGGCAATG GACCCCTTGG AACTGGAGGT GTGCAATGTT GTAAGCTACC
TTCAATTCAA 5220

GTCCTCTGGA AGAGCTGAAA GAGATCTTAA CAGCTGAGCC ATTCTCCAGA
CACCTGACCC 5280

TATTTCTTTG TCCTGGCGAG CACCTCCTGG GAAAAGTCTA CCTGGAGTAA
TGAGCAGACA 5340

TCTGACTCTT GCTCCTGATT GCTGGGAACA CATTCATAAT GGCACCATTG
AGAGTGCAGC 5400

TGACTGAAGT TGGACTTACT GCTTTGTTTT TAATGGATGA ATTTTGTAGG
GTTGAAGAAG 5460

TTCCTTTCAC TTCCATTTTC CACATATTTT TGTTGTTTAT TGAGATATTT
CTAATTTTTC 5520

CTCTTTTCTT ATTTTGTAAT ACATCTAATT ACATCAATTT ATTTTCTAGT
GTTTCTACAA 5580

CTCCTTCTAG CCAGGGCGCT TTTTATGGAG CCAAAACCAG ATTTTTGTTG
TTTGCATGTC 5640

ACTGGGATCT CCACTCCGTC CATTTTTGCT CTTCCATTAT TCACCCTGAG
TTCAGTGATC 5700

AGAAGTCTCC CTGGCAGAAG TTCTGTTTCC ATCCTCTCTA GACTTCTTTC
TACCATGGAC 5760

ATCCTTTGTG GACCAGGGCC TTCTCTAAAA GGGTTGGCAG AAGCCTTTCA
GGAACTGATA 5820

GGCAATTGTC AAGTGGGTTT TGTGGTTATT TTATTTTATT TTATTTAAGA
AATTCTTGAA 5880

ATTCGCATTC TTATATTCAT ATTTTACAGA TAAAATTCTC CAAAGAAAAA
GTCTCAGAGC 5940

CATCTCCCCT GAGTCTTCTG CTAAGCTTTA CTGCTGCTAT GGAGTGATCA
TGGTCCTCAC 6000

TGTAGCTGTA GTTGCTCTTT CTGTTGCTTT GTCAGGTAAG TGACATACCC TCCAAATTCT 6060

GTGACACTCT GTCCATATTC ACATTGCCAG TTATGCTTTC TAAGCACTGT GATCCAGGCA 6120

CTGTGGCAAG GGCTCTAGAG GAAACACACT GGAAGGTCCT GTTCTCTGAG AATTTAGGTT 6180

CCAACAGGAA GATGCAGTGA AGGAACACAG AGGCTTTGAT GGGGACATCC CCGGGAAGAT 6240

GACATCCAGC AAGCTCTAGA CAGAGATGCA GGGGACATAG GTCCCCTTGG GGGAACAATT 6300

CAAGGCAGAG AATAACAAGA GGGAATCTCC AAGTAGAAAC TTCAAAGGTG AGGCCAGGAA 6360

GGTACAGTGC CTTTGACCAT GACCCATGAG TTTAGATACC AGGCTCAACT CTATTTTGAA 6420

AGTATTAAAT GGAAAGTTCC TGAAGTAAGA AATTTATAGG ATTTTAGTAC CACAATATTC 6480

AGAATAGTGC AATACAATCT TGCACTGTCC TCTTAAGTAT TTGAAGTCAT CCTTTAGTGC 6540

AAGGTGTCTG CACCGTATAT ACGACCTACC CAAAAATTCT CATAGAAATC TCAATTATCA 6600

GGCTGGGTGT CAGTAGGTGT CCCTACAGAG TGCTTGCTGC TGTAGCAATC CCCACTGTAG 6660

TCAATGGTCA TCCAAAGCTC AGAAAGTGAT GCTGTTATAG AAGTGCACTC CCTGGGAGCC 6720

CTACTGACAG TGAGCACCTG AGAGAGAATG GGACACAGGC CCACGGTGGG
AGGCCTTTAG 6780

TTAAAGGCAC ATCTCGATCA GGAGAGGATT CCTACAGATC AGTTAGGAAA
GCTACCATCA 6840

GATTCACACC TCACAGCTGA GCTCAGGAGA GTGTGGCAAA ACGAGAGAAG
ACCTGCTTGC 6900

TATGATCCAT CATATTCTCT ACATTTTAGT AACAAAGACA GAACAGATCC
TAATCAACAA 6960

GACCTATGCT GCTTGCCCGA AAAACTGGAT TGGAGTTGGA AATAAATGTT
TTTATTTTTC 7020

TGAATACACA AGTAACTGGA CATTTGCCCA GACCTTCTGC ATGGCACAAG
AGGCCCAACT 7080

AGCTCGGTTT GACAACGAGA AGGAGCTGGT AAGCAATGGG CAGGGATTGG
TTTGTCTGTC 7140

TGTTCTGTTG AATATTATAT TGCCTTGAGA TAGAGAGTTA CAGATGAGGC
CCGAGGAAGG 7200

GATCCCACCC AAGCACATGG AGACATAGGG AATGTGAGTG TGTGCCATTT
GCTGATGCTT 7260

GACTTCTGAC TGGAGCCCTG AGATAGTCAA GAAACATTCT CTCATGAAGT
GCTCATAGTC 7320

AGCTGGAAGG TCAAATATGC CATTTTACTG GGATACCTGG TGACCATGAG
TGTTTTCCCA 7380

TATGCTGGCA TATGTTGGGT ACAGAAGGAG ACAACTGATA ATAACTGCAG
TGGAAGGTTA 7440

ACCCAGAACT GTCCAAACCA CAGAGGAATG TGACCCTCAG TTACATCCTC CTGTTATCTC 7500

TAGAGAAAGG TGTGGAGTGG AGAGACTCCA GGATCATCTG AAACAAATAG ACACATGTAT 7560

TCTTGACTTT TTTGTGTTTT ATGACAGAAT TTCCTAATGA GATACAAGGC AAATTTTGAT 7620

TCCTGGATTG GACTGCACAG AGAGTCGTCA GAGCACCCTT GGAAGTGGAC AGACAACACT 7680

GAGTATAACA ACATGTATGT TTTCACGATG TTTTTCCTTC TATTATGTTC ATGTGTTGTG 7740

ATATGTGTGT GTCGTGGCTA TGAGAGATGG AAGTCAATGT CATGTGAAGC CAACTGTACT 7800

GGGAAGAAAG AAAAAAAAAT GAACCCTTGC CTGGAGGTGT GGCTCAGGGG TAGAGAGTGT 7860

GTATAAATGC AATATCCAAT CCCCAGAAAG CTCTACACAC CCACAAATTT AAATACTTCA 7920

GAGGTTTTCC TGTTTATTGA CCGTCATTTT CAAAACCTTT GCATCATGTC ATTTTACTCA 7980

AAATATTTAC CATAATGATG GTGTCTGAGA GTAGCTATTG TTTGCTCTGG CTCCAACTTA 8040

AACATTTCTG TTGTTGATAA ATGTCCTGTG AGGGATATAG ACAGAGCCTT AGATGGGCAG 8100

TGGGGGCTCT GGAATCCCAG AAAGCCACTG CAGTATCTGC AAGCCTGAGA TTCAGCTTTC 8160

CACTATTTGC ATGTCTGCAC CTGTTCAGGA AAGCAGAGAC TCTAAGTACA
TTTGGAACCT 8220

CCTCTAAAGT CTCGTCATCA CTGAGCACCC AAAACAGTCT TGGGTTTGAG
CTGTTTTACT 8280

GGGATGGTAA ATCACAGACT CAGTCACATC CATCACTGAA GCCCTTAGAG
CAATTTACTA 8340

AGTGGGCGTC CCCATATATA AAATGCCTAA AACAGAATTG AAAATCACCC
TTGGTGGGGT 8400

CACTCATGGC TGCAGTTCAT TTGAACATGG CAGCGAGCAC CAGCCCAATG
CCTTGTACAC 8460

ACATTACAGG ATTCACCATG GACAAATGAC AAAGGAGTGG TGTTCAAATC
CTGAGAATAT 8520

GAGACAGTAG GTGTAAAACT AATGCAGGTG ATTCCTCAGG GACTTTTTGA
TTCATATTAC 8580

CAAAAATTAG TGGAGACTGG TGAGATTTCA TTGCAGGAGC AAATGCAGTT
CTGGGCTCTG 8640

TAGGCTTACT TTTTTGGTTT CTTTTCAGGA TTCCCATCCA GGGAGTGGAA
ACATGTGCCT 8700

ACCTGAGCGG CAATGGGATC AGCAGTTCCA GGCACTATAT ACCTCGGATA
TGGATCTGTA 8760

GCAAGCTTAA CAACTATAGC CTCCACTGCC CAACTCCTGT TCCTGTCTAG
CATTTACCAA 8820

GAGACTCTTC CTAGCCTGTT ATCTATGGGT GCTACTTTTT CCCCTATGGT
CCCACAGTGC 8880

TATCAAACGG GATTGAGAAT ATTTTTTAAC GTCGCAAATG AAAACCATCA AGGCTGGAGA 8940

GATTGCTCCG TAGTTAAGAG ACTGACTGCT CTTCTGCATG TCCCGAGTTC ACATCTGAGC 9000

AACCACATGG TGTCTTACAA ACATCTGTAA TGACATCTTA TGTCCTCTTC TGTGGTGTGT 9060

GAAAACAGCT ACACTATACC TACATATGAT AAATAAGTAA ATCTTAAAAA AGAAAAAGAA 9120

AACCACCTTA GAGAGGTGCA CACATGGAGG ATTACAAGAC CATAGATGAG TTTTAAATAG 9180

ATGTCAGCAC TCATACCTTA AGCCTAAAGT ACAACTAATG TTAGGGAACC CCACTTTTAT 9240

GATATTAAGG TTTTGTGCAG AGAATTCTTC TTTTGAATTT ATGAGACCAC AAAAATGAGT 9300

CCCCCAACAT GGGTGTAACC TTTAATAATG AAAGCAGAAT GGCTGGGAT 9349

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 990 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GATAGTGGTG CAGAGCCTCC CATGCCAGAT TGCTTGGAGA CAGGAGAAAA ACTGTTTGTA 60

CATAACATGA ATGCCCAGTG TGTACAGAAG CCAGAAGAGG GCAATGGACC CCTTGGAACT 120

GGAGATAAAA TTCTCCAAAG AAAAAGTCTC AGAGCCATCT CCCCTGAGTC TTCTGCTAAG 180

CTTTACTGCT GCTATGGAGT GATCATGGTC CTCACTGTAG CTGTAGTTGC TCTTTCTGTT 240

GCTTTGTCAG TAACAAAGAC AGAACAGATC CTAATCAACA AGACCTATGC TGCTTGCCCG 300

AAAAACTGGA TTGGAGTTGG AAATAAATGT TTTTATTTTT CTGAATACAC AAGTAACTGG 360

ACATTTGCCC AGACCTTCTG CATGGCACAA GAGGCCCAAC TAGCTCGGTT TGACAACGAG 420

AAGGAGCTGA ATTTCCTAAT GAGATACAAG GCAAATTTTG ATTCCTGGAT TGGACTGCAC 480

AGAGAGTCGT CAGAGCACCC TTGGAAGTGG ACAGACAACA CTGAGTATAA CAACATGATT 540

CCCATCCAGG GAGTGGAAAC ATGTGCCTAC CTGAGCGGCA ATGGGATCAG CAGTTCCAGG 600

CACTATATAC CTCGGATATG GATCTGTAGC AAGCTTAACA ACTATAGCCT CCACTGCCCA 660

ACTCCTGTTC CTGTCTAGCA TTTACCAAGA GACTCTTCCT AGCCTGTTAT CTATGGGTGC 720

TACTTTTTCC CCTATGGTCC CACAGTGCTA TCAAACGGGA TTGAGAATAT TTTTTAACGT 780

CGCAAATGAA AACCATCAAG GCTGGAGAGA TTGCTCCGTA GTTAAGAGAC TGACTGCTCT 840

TCTGCATGTC CCGAGTTCAC ATCTGAGCAA CCACATGGTG TCTTACAAAC ATCTGTAATG 900

ACATCTTATG TCCTCTTCTG TGGTGTGTGA AAACAGCTAC ACTATACCTA CATATGATAA 960

ATAAGTAAAT CTTAAAAAAA AAAAAAAAAA 990

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TCCCATGCCA GATTGCTTG 19

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GGGACCATAG GGGAAAAAGT AG 22

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 721 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

TCCCATGCCA GATTGCTTGG AGACAGGAGA AAAACTGTTT GTACATAACA TGAATGCCCA 60

GTGTGTACAG AAGCCAGAAG AGGGCAATGG ACCCCTTGGA ACTGGAGGTA AAATTGTCCA 120

AGGAAAATGT TTCAGAATCA TCTCCACTGT GTCTCCTGTT AAACTTTACT GCTGCTATGG 180

AGTGATCATG GTCCTCACTG TAGCTGTAAT TGCTCTTTCT GTTGCTTTGT CAACAAAAAA 240

GACAGAACAG ATCATAATCA ACAAGACCTA TGCTGCTTGC TCAAAAAACT GGACTGGAGT 300

TGGAAATAAA TGTTTTTATT TTTCTGGATA CCCACGTAAC TGGACATTTG CCCAGGCCTT 360

CTGCATGGCA CAAGAGGCCC AACTAGCTCG GTTTGACAAC GAGGAGGAGC TGATTTTCCT 420

AAAGAGATTC AAGGGGGATT TTGATTGCTG GATTGGCCTG CACAGAGAGT CGTCAGAGCA 480

CCCTTGGAAG TGGACAAACA ACACTGAGTA TAACAACATG AATCCCATCC TAGGAGTGGG 540

AAGATATGCC TACCTGAGCA GCGATAGGAT CAGCAGTTCG AGGAGCTATA TAAATCGGAT 600

GTGGATCTGT AGCAAGCTCA ACAACTATAA CCTTCATTGC CAAACTCCTC CTGTCTAGCA 660

CTTACCAAGA GACTCTTCTT AGCCTGTTAT CTATGGGTGC TACTTTTTCC CCTATGGTCC 720

C 721

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

TGGAAACTCA GCTCCTCAGC TCTG 24

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 713 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

TGGAAACTCA GCTCCTCAGC TCTGAGATGT GTGTCACAAA GGCTTCCCTA CCTATGCTTA 60

GTCCCACAGG CAGCCCGCAG GAGGTAGAAG TGGGTAAAAT TCTCCAAGGA AAAAGGCACG 120

GAACCATCTC CCCTGAGTCT TGTGCTAAGC TTTACTGCTA CTATGGAGTG ATCATGGTCC 180

TCACTGTAGC TGTAATTGCT CTTTCTGTTG CTTTGTCAGC AACAAAGACA GAACAGATCC 240

CAGTCAACAA GACCTATGCT GCTTGCCCGC AAAACTGGAT TGGAGTTGAA AATAAATGTT 300

TTTATTTTTC TGAATACCCA AGTAACTGGA CATTCGCCCA GGCCTTCTGC ATGGCACAAG 360

AGGCCCAACT AGCTCGGTTT GACAACCAGG ATGAGCTGAA TTTCCTAATG AGATACAAGG 420

CGAATTTTGA TTCCTGGATT GGCCTGCACA GAGAGTCGTC AGAGCACCCT TGGAAGTGGA 480

CAGACAACAC TGAGTATAAC AACACGATTC CCATCCGGGG AGAGGAAAGA TTTGCCTACC 540

TGAACAACAA CGGGATCAGC AGTACCAGGA TCTATTCACT TCGGATGTGG ATCTGTAGCA 600

AGCTCAACAG CTATAGCCTC CACTGCCAAA CTCCTTTTTT TCCTTCCTAG CATTTACCAA 660

GAGACGCTTT TTAGCCTGTT ATCTGTGGGT GCTACTTTTT CCCCTATGGT CCC 713

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

TTTGTCAGCA ACAAAGACAG AACAG 25

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1229 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

CGGGTGGGCG CGAGAGAGCC TAGAAGCCCA TGTAGCCGCG AATCCCGCAG CCCCAGTACA 60

CCTCCCTCCG TGCCTCCCCG CCTTTTCTGC AGAGCTCCGC CCTGGAGTGA AGGAGGAGCC 120

GTCACCTGGA GCTCCGAAAA AAGCAGAAGA AGGCGCTTTT TATTTAGCCA GTGTGACCCC 180

GCCAGGGCCT TCTCGGTTGG GTGAGCACTC TCTCTGACCA GGCCATGAAA AGAAAAATCT 240

GTGCGATGCC TCCCCACATG TCACGGGACT CTGACTTGCC TTTGTCGTCA
GAGTTTGCAG 300

AACTTTGGGG GACCTGAGAG GGGAGTGCCC CCTGGACGGG CCACGGCTGT
CTGTGGCTTA 360

AGGGCTTTTG GAAGGGCGGA GAGAGGGAAA CGGCGTCCTA GTGGCCTGCT
TCAGGGCCAC 420

CCACGGGCCC TCCCCCAACC TCTCTCTGAT CCAACTTGTT TTTCCAGCCT
AGTTGGAAAC 480

TTGTGGATGC TGTGACCTCA AGAAGACTTG GCATTTTATT TGGAAGATAG
ACATCTATTT 540

GCAACTGGTC CTGAGCCCCT ATTTTCCTCC CACCTTTCTT GGGGAAACTT
GTTTTTAAGG 600

GGTGCCACTG TTTTTGTAAC ATGTTGCTCC TAGCTCTTAG CATTCATGGT
ACTGTTGTAA 660

ATGGAGAAAG AGTAATTCAC GCAGAGCCGG CTTTGCAGAT AAAACTCTGC
AAAGACAATG 720

TCTAGCACTT ATATCTCCAG TAACTCCTGT CAAGGTTTAT TGTTGCTTGT
TCATCCATTG 780

CAGTCCTGAC TACAAGTGTG ATTGCACTTT CTATTGTTTT GTCAGAATTT
CCTGAAAAGA 840

TACAAAGGCC CTTCTGACCA TTGGATTGGC CTTAGAAGAG AATCATCCCA
TCGCATTTGG 900

AAATGGACAG ACAACATGGA ATATAATAAC ATGCTTGCTA TCAGAGGAAG
TGGAGAATGT 960

GCCTTCCTGA ATGACAATGG AGTCAACAGT GGCAGAATCT ACATGAACAG AAAATGGATT 1020

TGTAGCAAGC CAAACAATTA TGTCTACAGT TGCCAGTTAT GTCCCCACTG GGATACTACC 1080

TAGTAGAGCT GTGAGAAGAG GGCCACCATC CTCCAGACTC CAGAATGGTG GAATCATCAG 1140

CAGCTTCCAC CATGCCCCTG GAAAAACTGC AAGTAACAGA CCTGCACATG TATCCCCTAC 1200

ATCTAAAAAA AAAAAAAAAA AAAAAAAAA 1229

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1336 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

CGGGACAATG TTATGTGGCT CAGAGGCCCT CCATGTATTC TTAACTATTC ACTCTCCTAT 60

CCTTCCAAGA ATAACACTAA TTTGACCTCT ACAATAATCA CTTTATCACT CCAGTTTTGC 120

CTTTTTTCCT CCAAAACAAT GCCTTTTAAG TCTATTTTAA TCGATAGATT TCCTCTTAAT 180

ATCATTTAAA AATATTTCTT TACATTTTTA GACAGGAATC AGAATAATTT GCTATGTTGA 240

ATTTCCAGTT ACTTGGATTT TGTTGATTTC ATTCCTGTGG TTTAGTTGAC ATGAATCTCT 300

CCAATTGAAA GGGTAACTTG AATATGGTAG CTGGAAAGTT AAAATCAATT CTTTTAACTT 360

TGGAAAATGA TTAAATTCTG GAGATAGAAT AGGTAAGGTT CATAAGATGA CAGGGTCGTT 420

TGCATCCTTC TAGTGGAAAA GCGAAGGGAA TTAAATAAAA ATAACACTTT GATGCTTAAT 480

GTTTCTGGCA GTATTATGTC TGTATTTAAT TGTTAAAATG TTTTTCAATA ATTTTTTCCA 540

AGCGGGTTGT CTGCATTCAA AAGAGCATTC TATTAAAGCT ACCTTAATTT GGCGCTTATT 600

TTTCTTAATC ATGTTTCTGA CAATCATAGT GTGTGGAATG GTTGCTGCTT TAAGTGCAAT 660

AAGAGCTAAC TGCCATCAAG AGCCATCAGT ATGTCTTCAA GCTGCATGCC CAGAAAGCTG 720

GATTGGTTTT CAAAGAAAGT GTTTCTATTT TTCTGATGAC ACCAAGAACT GGACATCAAG 780

TCAGAGGTTT TGTGACTCAC AAGATGCTGA TCTTGCTCAG GTTGAAAGCT TCCAGGAACT 840

GAATTTCCTG TTGAGATATA AAGGCCCATC TGATCACTGG ATTGGGCTGA GCAGAGAACA 900

AGGCCAACCA TGGAAATGGA TAAATGGTAC TGAATGGACA AGACAGTTTC CTATCCTGGG 960

AGCAGGAGAG TGTGCCTATT TGAATGACAA AGGTGCCAGT AGTGCCAGGC ACTACACAGA 1020

GAGGAAGTGG ATTTGTTCCA AATCAGATAT ACATGTCTAG ATGTTACAGC AAAGCCCCAA 1080

CTAATCTTTA GAAGCATATT GGAACTGATA ACTCCATTTT AAAATGAGCA AAGAATTTAT 1140

TTCTTATACC AACAGGTATA TGAAAATATG CTCAATATCA CTAATAACTG GGAAAATACA 1200

AATCAAAATC ATAGTAAAAT ATTACCTGTT TTCATGGTGC TAATATTACC TGTTCTCCCA 1260

CTGCTAATGA CATACCCGAG ACTGAGTAAT TTATAAATAA AAGAGATTTA ATTGAAAAAA 1320

AAAAAAAAAA AAAAAA 1336

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10221 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GAATTCCTTC TTTTTCTATT GTTTGGAATA ATTTCAGAAG GAATGGTACC AGCTCCTCTT 60

CGTACCTCTG GTAGAATTCG GCAGTGCATT TGTCTGGACA TGGGCTTGTT TTGGTTGGGT 120

AGGCTATTAA TTACTGCCTC AGTTTCAGAA CCTGTTATTG GTCTATTCAG
GAATCTGATT 180

TCTTCCTGGT TTAGTCTTGG GAGGGTGTAT GTGTCCAGGA ATTTATCCAT
TTCTTCCTTG 240

CCTGGGTATC ACCAGCAAAG GCTGAAGAAA AGCAAAGATT GCTGCCTGCT
CCTTCCTCTG 300

GAAGCTTCAT CCCAGAGGGG CACGCACCAG ATGCCAGCTG AGCTGTCCTG
TATGAGGTGC 360

CTATCAACCC CTGCTAGGAG TTGTCTCCCA GTCAGGAGGC ATGGGGGTCA
GGGACCCACT 420

TGAAAAGGCA GTCTTTCCCT CAGAAGAGCT CGAGCACTGT GCTGGGAGAT
CCACTGCTCT 480

TTTCAGAGCT GGCTGGCAGG AATGTTTAAG TCTCCTGAAG CCGTGACCAC
AGCCACCCTT 540

TCCCCCAGGT GCTCTGTCCC AGGGAGATAA GAGTTTTATC TATAAGCCCC
TGACTGGGGC 600

TGCTGCCTTT CTTTCAGAGA TGCCCTGCCC AGAGAGGAGG AATCTAGAGA
GGCAGTCCGG 660

CTGCAGTGGC TTTGCTGCAC TGGCTTTGCT GCACTGTGGT GGGCTCCGCC
CAGTCCGAAC 720

TTCCCCCAGG GCTTTGTTTA CACTGTGAGG GGAAAATCAC CTACTCAAGC
CTCAGTAATG 780

GCGGATGCAC CTCCCCTCAC CAAGCTTGAG CATCTGGGGT CCACTTCAGA
CTGCTGTGCT 840

GGCAGCAAGA ATTTCCAGCC AGTGGGTCTT AGCTTGCTGG GCTCTGTGGG
GATTGGACCC 900

ACTGAGCAAG ACCACTTGGC TCCCTGGCTT CAGCCCCCTT TCCAGCAGAG
TGAATGATTC 960

TGTCTCAGTG GGTTCCAGGC TCCACTGGGG TATGAAAAAA ACTCCTGCAG
TTATCTTGGT 1020

GACTGCCCAA ATCGCCACCC AGTTTTGTGC TTGAAACCCA GGGTTCTGGT
AGTGTTGGCA 1080

CTCCAGAGAA TCTCCTGGTC TGTGGGTTGC AAAAACCGTG GGAAAAGCGT
AGTATCTGGG 1140

CCAGATAGCA CCTCACAGCA CAGTCCCTCA CAACTTCCCT TGGCTAGGGG
AGGGAGTTCT 1200

CCCACCCCTT GTGCTTCCTG GGTGAAGCAG CGCCCCACCC TGCTTCTGCT
TGCCCTCTGT 1260

GGGCTGCACC CACTTGTGTA ACCAGTCCCA GTGAGATGAT CCTGGTACCT
CAGTTGGAAA 1320

TGCAGAAATC ACCTGCCTTC TGCATTGGTC TCACTGGGAA CTGCAGACCA
GAGCTGTTTC 1380

TAATCAGCCA TCTTGCCCTC TCTGGTCTGG TCGTTTTCTT TAAATTGGTT
GATACAGGAG 1440

CAGTGATAGC ACAACAAATA TGCACAGATT TGGGGAAAGT CATCCTGCAT
TATGGTCTGG 1500

TTCAAGAAAT TACATTTTAA TAGTTATAAT TTGGTATCAC CTTGTTTGTG
ATACCAAACC 1560

AGATACAATA CACGTTTGCC TCATGTTATG ATTTGTTATT CAGATTACAC CAGTTATTAT 1620

TCATAACTAA GAGTGATTTC TCATCTCACA AGAGCCAAAT CCAAGGATAA TGGTGCCAAT 1680

TGATAGTAAT GATTCTATGA ATACCCAGCA TTCTGGTCTA TCATAGACAC TTTCAGAACC 1740

ATTGAGTTGA AGGTAGAAGG TGGTTATATA ATAGAAGATG AACTGGTAGC TACTAGGGGC 1800

TCAGTGACAA CTATCTGGAG AGACATTCAT TCATCTCGAT CCCACATGAA GAGAGCATTT 1860

CTCCTGATTA TATAAGAAGT GGGTCAGAAA AGCCTGTCCA GTGAAGTATT GCTGCCTTCA 1920

AAGTGTAGAA AACCTCACTA AATCTCCTTA GTGGAAGGAA GTTCACTGTA CAACAACTTA 1980

TTTCATATTT ATGATAGTAT TTAGACATAT ACAAGGCTTT TTCACATCAA GAAACCTTAT 2040

TCACATAAGG CATCTCTATC CTGCCCTTCA TTTTACCAAG TCATCTGGAG CAGCAATCGC 2100

CAACCTTGTT GGCATGAGGG ACCAGTTTTG TGAAAGACAA CGTTTTCATG GACTGGGGTC 2160

AAGGAATGGT TTGGGGATAA TTTAAGTGCA TTACTTTTAT TGTGCCCTTT ATTTCCATTA 2220

TTATTACATT GTGTAATAAT ATATAATAAA ATAATTATAC AACTCACCAT AATGTAGAAT 2280

CAGTGGGAAC CCTGAGCTAG TTTTTCTGAA AGTAGATGGT ACCATCTGTG AGTGATGGGA 2340

GACAGTGACA GTTCATCAGG TATTAGATTC TCACAAGGAG CCCACAACCT AGATTCCTCA 2400

CATGAGAAGT TCCCAATAGG GTTTGCCCTC CTATGAGAAT CTAATGCCAC TGCTGATCTG 2460

ACAGGAGGTG GAGCTCATGA GGTAATGTGA GTGATGGGGA GTGGCTGTAA ATACAGATGA 2520

AGCTTCACTT ACTCATTCGC TGCTTACCTC CTGCTGTGCA GCCTGCTTCC TGACTCATCC 2580

ATGGACCAGT ACTGATCCAT GGCCTAGGGG TTGGGGACCC CTAATCTAGA GCACTTGGAG 2640

AACTATCTGT TCTCCAAAGC TGATCAAATG CTATCATTAA TGTATCTAAT ATTTTAAGAA 2700

AGGGTAACAC TGTTGAGAGC CAAATAGATA CATGGCCCAG AGCAAGCTTA AGTTACTAAT 2760

AACTCCTTTT TCAGCTCACC CCCTGCTGAA GGCATGAGTT TGAATCTCAG TTTTGCCATT 2820

TGCTGTGTAA TGTATGCAAT TATATTTAGC ATCATATTTC TCACTTGAAA AATGAAAATA 2880

ATACATTTAA TACTTAACAG GAGTGTCAGA AAGTATATTA GCACTTGGTA ATTTATACAA 2940

TACAATATAA AAGTAAGAAA TTTTTATTTT ATTTATTTTT ATTTTATTTT CAGCAATAAG 3000

AGCTAACTGC CATCAAGAGC CATCAGTATG TCTTCAAGCT GCATGCCCAG AAAGCTGGAT 3060

TGGTTTTCAA AGAAAGTGTT TCTATTTTTC TGATGACACC AAGAACTGGA CATCAAGTCA 3120

GAGGTTTTGT GACTCACAAG ATGCTGATCT TGCTCAGGTT GAAAGCTTCC AGGAACTGGT 3180

AAGAAAATAG TTCTGGCCAG AATCAAAGAT TCAGCCCTAC AAGGATATGT TTTCCTGTGA 3240

AATTATCTAA GAGGTAGGTT TAGACATCTG CTTTTACATT GATTTTTTTT TTTTTTTTTT 3300

TTTTTTTGCAT AACGAAAGAG TAACCTAGCA TGTATTATAT TTTACAGTGA ACCATCTAAA 3360

ATTACCTTAA TATTCGTGGC AGGAACAGGC CCAGAGGGCA AGCAAGCCAG AGCCTTCTTT 3420

GACTTGTGAG CCAGAATTGT GCAAATAAGG ATTAGAAAAG TATTGGTAGA AACCCAGTTT 3480

TAAGTTTGTA TGAAGTTAGC AACATTGTTT CAAAATAAAT CAAACAAGGC CAAGAGCAGT 3540

GGCACATGCC TGTAATCCCA GCACTTTGGG AGGCCAAGGC GGGTGTATCA CTTGAGGTCA 3600

GGAGTTTGAG ATCAGCCTGG CCAACATGGT GAAACCCCAT CTCAACTAAA AAATACAAAA 3660

ATTAGCTGGG CATGGTGGCA TACGCCTGTA GTTCCAGCTA CTCAGGAGGC TGAGGCAGCA 3720

GAATTGCTTG AACCTGGGAG GTGGAGGCCT ACAGTTAGCT GAAATCATGC TACTGTACTC 3780

CAGCCTAACA GAGTGAGACT CTATCTCAAA AAAATAATAA AATAAAAACA ATAAGTCAAG 3840

CAAGAATGAT GTCATAGAGG TTGGTAGACT AAAAAGCTAC AGAAATCTGT TCCTCCACTG 3900

AGAAAACTAT TGAACTGTCA AAAACTGTCT GAAGTAACTA TTTTGGAATT CTCGAGTCTA 3960

GTTAAACACT GGAAGCATCA AGGGAAGAGT TTGATAAAGA GGATGATAAA TTTTGGTTAA 4020

TGTTGGTGAA TTTCAGCCTT TCCACTCAAT AATAACTATT TTCCATACCC CATTATTGCA 4080

GGGATCCATG GGAACTGCTG CCCATGTTCT TGTAATGAAT TCCTGCAGCC AGGGTGAACA 4140

ATAAGCACCT TTTTGTCCAA ATGTCAGGGT TATTGCTGAT TTCTGCCTTT GAATGCTGAG 4200

GGGCAGACAC AGAAGTGGGC TATCATTGCA TCAGTCCTCA TCAGCTGAAG TGGCTTCCCA 4260

AGGATTTAAA TAAATAGTAT GTGTTTTTCC TCCCTTTAGG AAGCAGTCAT TTAAGACAAT 4320

TTTTATTAGA TAACTGGCTG ACAGCAGAGA TAACAGAACA GAGATTTCAA TGACCATGCA 4380

CAACAGAGAA TAAAAATAGT TGGGAAAAAA TCATGACCAA ATGACTCTGA GCCACAACAA 4440

CCAAGATTTG ACAATCCCTG AAGAGCAAAA TAATTAAGTT ACCAGAGTTA CCACAACATA 4500

GTATTCATAA TGTCCAGTTC TCAAAAAAAA ATTACAAAAC ATGCAAAGAA AAGTATGGTT 4560

CATTCACAGG AAGAAAAAGT AATCTGACAG AAACTATCCC TGAAGAGGCT CAGATATTAA 4620

AAATATGAGT CAAAAATGTT AAATCAGCTG TCTTAAGTAT AACCAATGAG TTAAAGGAAA 4680

CTAGACAAAA AGCTAAAGGA AACCGAAAAC ATAATAAATG AACAAAATTA GAATATCAAT 4740

ATAAAGGTAG AAATTGTAAA AAAGAACCAA GCAAAAATTC CAGAGCTGAA AAGTACAATG 4800

ACTGAAATTT AAAAATAATT TTAAAAACTC AATGAAGAAG TTCAACAGCA GATTTGAGAA 4860

GTAAGAGATC AGAAAACTTG AAAATAAGAT AATTGAAACA ATCCAGACTA AGAAAAACAA 4920

AGAAAAAGAA TGAAGATAAA TAAATTCTAA GGAACCTGTA GGACATCAGC AAACATACTA 4980

ACATATGTAC TGTAGAAATC CAGGAAAGAG AAGAGAAAGA GAAGCAGAGA AATACACTTA 5040

AAGAAATAAT GAACAAAACT TTCCAAAATC TGAGGAAATA CATAAATATA TACATCCAAG 5100

AGGCTCAATG AACTCCAAAA GGGTAAACTT AAAGAGATCT ACATTGAGAC AAAATATAGT 5160

CAAGTTGACA AAATCCACAG AGAGAATTTT GAAAGCAGCC AGAATGAAGC AACTCATCAT 5220

TTACATAAGA CCCTGAATAA AATTAATAGC TGATTTTCTC TGAGAAACCA TGGAGATCAG 5280

AAGGTAGTGG AATGGCATAT TTAAATGTCT GAAAGAAAAA ATAAAACTGC CAACCATGAA 5340

TTCTATGTAT AGCAAAGTTG TCCTTCAAGA ATGAAGGAAA AAGTAACACA TTTTCAGATA 5400

ACCAATAATT AAGGGATTTT ATTACCAGTA GACATGTGCT ACAGAAAATG CTAAAGGAAA 5460

CCTTTTAGGC TGAACTGAAA GTACACTAGA CAGCAATTCA GAGCCTCCAA AATAAAGAAT 5520

ATTCATAAAA GTAACAATAG AGGTAAATAT AAAACCCAGA ATTACTACAT GTGTCATATA 5580

GTTTATAACT TCTCCTATTT ATAGCTTTCT ATATTTATAT TTATCTATAA CTTCATAGGC 5640

AAATGAATAA AAATTATAAA TATGATAGTG GTCATATAAT GTATAAAGAT GCAATCTGTG 5700

ACAGTCTTAT GAAGCAGGGA TGAAGACATA TAGGATCAAA ATGTTTGCAT AGTTATTGAA 5760

GCTATGTTGA TATTATGAAA TTATATTGTT ACAAGTTTAA GATGCTAATT ATAATTCTCA 5820

AGGTAACCAC TAATAAAATT ACCAAAATTA TGCAGAAAAG GAAAAAAGAA AAACAATACA 5880

CTATAAAAAA CCAATTAAAT ACAAAAAAAG TCAGTAACAG ACAACTTGAG
AAACAAAGAC 5940

ATATAAGATA TAGAGAAAAC AAATGATTAA ATGGCAAAAG TAAATCTTGT
TTTAGTAATC 6000

ACATTAAATA GAAAAGGATG AAGCCATCCT ATTAAAGGGC TGAGACTGAC
AAGTTGGCTA 6060

AAAACTAAAA TAAATTAAAA AGAAAAACAA GACTCATCTA CATGCTGTCT
ATAAGAGACT 6120

TGCCTTAGAT ATAAGGACAC AAAGAAGTTG AAAGTAAAAG GACTGAAAAA
GATATTCCAT 6180

ACAAACAGTA GTAACCAAGA TAGTGCCGAG TGGCTATATT TTTGTCAAAC
AAAATAAACT 6240

AAAGTAAAAT TTACAAGAGA AAAAGAAGGG CATTATGCAT TGACAAAAAT
TTTGACATAG 6300

CCAAATAATT ATGTTATAAA ATATATGTAC TTAATAATAC AGCCTCAAAA
TATATGAAGC 6360

AATAATTGCT ATAATTTAAG GGAGAAAAGA ACAGTTCTAT GAAAAGTTAG
AGAATGAAAT 6420

ATTCCACTTT CAACATGAGA TTAAACAACT AGACATAAGA TCAATAAGGA
AATAGAAAAT 6480

TTGAACAACA CTATAAACCA ATTATCCCTA ACAGGCATAT ACAGAAGAAT
CTACCCAACA 6540

AGAGCAGAAT ATTAATTCTT CTCAAATGCA CATGGAACAT TCTTAAACCA
TATGTTAGGC 6600

CACAAAACAA GTGTTAGTAA GTGTGAAAAT TTGAAGTCAT AAAAAGTATC TTTTGCAATT 6660

ACAATGGAAT GAAGCTAGAA ATCAATAACT AGAAAAACCA GAAAAGTCAC GCATATGTAG 6720

AAATTTAAAA ACCCGCTCTT CAACAGCCAT TGGTCAAAGA AGAAATCACA AGGGACATTA 6780

GAAAATACCT TGAGACAAAT GAAGTAAAAA TACAAATAGC ACGTTTATGG TATACACTGA 6840

ACATAGTTCT AAGAGGGAAA TTTATAGCTG TGAGCAGTTA ACTAAAAAAG AAGAAAGATC 6900

TCAAATCCAT AGCCTAACTG TACACTGTAA GGAACTAAAA AAAGTAAAAC AAAAATAGAA 6960

GTCATCTTTA TGATTTGAAA GAGTAAAAGA TTTACCTAAT AAGTCCCTAA ATTTACTAAT 7020

AATAAAGAAA ATTGTTTATA TATTTAATTG CGTTAAAATT CAGAACTTGT AATCATAAAA 7080

AGGACAGTAC ACATTGACAA GGAAACACAG CAAAGGAAAC CAGCCTATGC TGCTGCTGTT 7140

GTGAGGATAA TTTGGTACAC TTACATTAGT TTGGTGTCTT TTCTTTCTCT TTCTTTCTTT 7200

CTTTCTTTCT CTTTCTTTCG TTCGTTCGTT CGTTCGTTTC TTTTTGAGAC AGAATCTCAC 7260

TCTATTGCCC AGGCTGGAGT GCAGTGGCGT GATCTTGGCT CACTACAACT TTTGTCTCCC 7320

AGGTTCAAAT GATTCTCATG CCTCAGCCTC CCAAATAGCT GGGATTACAG GTGCATGCCA 7380

TCACGCCCAG CTAATTTTTG TATTTTTTTT AATAGAGAGG GGGCTTCATC ATGTTGGCCA 7440

AGCCTAGTCT CAAACTCTTG GCCTCAGGTG ATCCGCCTGC CTCGGCCTCC CAAAGTACTG 7500

GGATTACAGG TGCTTGGCCT GGTGGTGTCA TTTCTTAAAG TTGACAAAAA GCATATCCTG 7560

GGGCCTAAAA ATTCTATTCT AGGACAGGTG CCAAGAATGT CATAGTAGCA TACATTCCAA 7620

ACTTGATAAA ACCCTGGTGT CAACCGATAG TATAATAGAT AAATTGGAGA AGAGTCATAC 7680

AAAGGAGTAC AATACAGAAA CAAAAGTAAC CAAATTATCA ACAATTTCTC TCAGTTTTAA 7740

ATTATCTTCT TTTGATATGT ATGATAATAT AGCACACCTA TTCTGTATGT ATTACTAAAC 7800

AATACAAAAT CAAAAGGAAG AAAATTATGA GTAGTTAAGA ATATAGCATA GCAGCAACAT 7860

TTCTGGGAGA GGATGGGTTA TGTTAGATTA ATGAATATCA TCTCTGTGTT TTCTGAAAGA 7920

ATTTCCTGTT GAGATATAAA GGCCCATCTG ATCACTGGAT TGGGCTGAGC AGAGAACAAG 7980

GCCAACCATG GAAATGGATA AATGGTACTG AATGGACAAG ACAGTAAGTT CTAAAAATCT 8040

GGCAGTAATA TTTGTATTTG AATTTACTTT GCATTAAATC TGAAGTGTTC TCTAGTTACA 8100

TGCTTTAAAA AATTCTCATT TTAAGGTTAG TCATGAAAGA AGATGGTGCC AACTTGTATG 8160

TTGCAAAGGT TTCACAAGTT CCTCGAATGA ATCCAAGACC TGTCATGGTG AGGTAGACTG 8220

ACTGTGAACT TGGCTCCAGG CTTATCTATG TCATTTTCAA ACACTTTCAT TTTAAGCAAA 8280

CCATACAATA TCTTTAAGTC TGTTCCTTAC CTCCACAACA AAATTAAATT GCACTTGTCC 8340

TCCTGATTTC ACAGGGTTGA TGTGAGGAAC AGAGGTTTTG ATGTATCAGG GAAAGATTAT 8400

GAGTGACAGC AATTATACCT ATTATTTAAA ATAAGACAAT AGTTTTAAAA TTTTAAAATG 8460

GGTAAAGTTT GGCACTAGAA AATTTAATCT CAATTGTATA TTTATAGGAT CTTCAGATTA 8520

CTAAAAAGAT TTGAGATAAT GCTGGAAAAA TTGGATTCAC ACAATTTCAC TCAATGTTTG 8580

TCTGAGAGAT GAGACAGTTT TGAAAAGCTA CTTTATTGTA ATACATTCAT CAATATTGGA 8640

AATATAACTT TATTTAATAA AAAGAGCCCC AGACTGGACA TTGGCAGGTT TGAAATGAGT 8700

TTTTTCTCAT TAGCTTTTGA CCTTGGATGG GATGGTAAGT TTTAGAAATC AGAGAACATG 8760

TACATTTATA CATTGTTGTA TCTACACTGC CTTGCACATT GTGACTGCTT CATAAATATC 8820

TAGAATTAAC TTTTATTTCT TATTTTACAA TACGGAAGTA GTAAATTTTC TCTACCTAAT 8880

TCTTAAAATG GTTTTTTGTT TGTTTGTATT TTTGAGAGAC AGGGTCTTAC TCTGTTACCC 8940

AGGCTGGAGT GCAGTAGTAC CATCGTGGCT CACTGCAGCC TTGACTTCCC TGGCTCAAGT 9000

GAGCCTCCCA TCTCAGCCTC CTGAGTAGCT GGGACTACAG GTGTGTGCCA CCTTGCTTGG 9060

CTTTTTTTTT TTTTTTTTTT TTTTTTTCAG CGATGGGGTC TCACTATGTT GCCTGGGCTG 9120

ATCTTGAACT CCTGAGCTCA AGCAATCCTC CCACCTCGGC CTCCCAAAAT GTTGGAATTA 9180

CAGGTGTGAG CCACCATGCC TGGCCTCTCA AAATATTTTA AGGATCAAAT ATATTATTAA 9240

CTAACCAGTT TTTGGAAACT GCTCATCACT TAAAGAAATG TAAAATATTA TATGATTAAG 9300

GTCTAACAAG TTTCAACAAT TAGCAAATTA TATCATAGAT GATAGTGATT CCAATGAGCA 9360

AAGAGGAAAA ATTTATAATC CAAATGCTGA CCTAAAATAT CTGTGCCAAG CCATCTAAAC 9420

TCAGCTAAAT AGCACTGCAG TTTCAGTACT AAAACCACCA GGGAAGTAGG AGGAATAAAA 9480

TCAAGCATGG TTTTTAGAAA TAGCTGCTGA GTCTTCAGTT ATTTAAGGAA GCAAAATATT 9540

GGGAAACTGT GTAAAGAAAA CGTGTCAGAC TTCTCCCATC AGCCAGCTAA GGCTTTGGAT 9600

GTACTTGAAA GAATATTATG CTTACAGACA TGAAATAGGT TTGATTCAGG ACTTTGCAGT 9660

ATTCCTATAG TTGATTTATA ACATCTCCTG CTAAGCAAAG CCCACTGACT AATTAGTCAC 9720

CACTACACAA GGAAAAACAG CATTATTTTT AGAGGCTGAA TTAATGTTAG TTTTCTCATT 9780

TTCTCATCTT CATTCTCTCT GCTGTTGAAG AAATGTTCAG TGGCCAACTG ATTCTGCTTC 9840

TTCTCTTGCA GGTTTCCTAT CCTGGGAGCA GGAGAGTGTG CCTATTTGAA TGACAAAGGT 9900

GCCAGTAGTG CCAGGCACTA CACAGAGAGG AAGTGGATTT GTTCCAAATC AGATATACAT 9960

GTCTAGATGT TACAGCAAAG CCCCAACTAA TCTTTAGAAG CATATTGGAA CTGATAACTC 10020

CATTTTAAAA TGAGCAAAGA ATTTATTTCT TATACCAACA GGTATATGAA AATATGCTCA 10080

ATATCACTAA TAACTGGGAA AATACAAATC AAAATCATAG TAAAATATTA CCTGTTTTCA 10140

TGGTGCTAAT ATTACCTGTT CTCCCACTGC TAATGACATA CCCGAGACTG AGTAATTTAT 10200

AAATAAAAGA GATTTAATTG A 10221

(2) INFORMATION FOR SEQ ID NO: 22:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: YES
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

GAGTGTTGTC TGTCCACTTC C 21

(2) INFORMATION FOR SEQ ID NO: 23:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: YES
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

TTTCCAACTC CAATCCAGTT T 21

(2) INFORMATION FOR SEQ ID NO: 24:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GAGGAGCTGA GTTTCCACTA C 21

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

GGT AGG GAA GCC TTT GTG AC-3'.

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 amino acid

(B) TYPE: amino acid

(C) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

H-Cys-Met-Arg-Thr-Glu-Ala-Gln-Leu-Ala-Arg-Phe-Asp-Asn-Gln-Asp-Glu-Leu-Asn-Phe-OH

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

GCCACGCGTT TGTCAGCAAC AAAGACAGAA CAG 33

(2) INFORMATION FOR SEQ ID NO: 28:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: YES
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

GCCACGCGTG GGACCATAGG GGAAAAAGTA G 31

(2) INFORMATION FOR SEQ ID NO: 29:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 660 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

ACAATGGTTC TTGCCAGCTC TACCACCAGC ATCCACACCA TGCTGCTCCT

GCTCCTGATG 60

CTCTTCCACC TGGGACTCCA AGCTTCAATC TCGGCGCGCC AGGACTACAA GGACGACGAT 120

GACAAGACGC GTTTGTCAGC AACAAAGACA GAACAGATCC CAGTCAACAA GACCTATGCT 180

GCTTGCCCGC AAAACTGGAT TGGAGTTGAA AATAAATGTT TTTATTTTTC TGAATACCCA 240

AGTAACTGGA CATTCGCCCA GGCCTTCTGC ATGGCACAAG AGGCCCAACT AGCTCGGTTT 300

GACAACCAGG ATGAGCTGAA TTTCCTAATG AGATACAAGG CGAATTTTGA TTCCTGGATT 360

GGCCTGCACA GAGAGTCGTC AGAGCACCCT TGGAAGTGGA CAGACAACAC TGAGTATAAC 420

AACACGATTC CCATCCGGGG AGAGGAAAGA TTTGCCTACC TGAACAACAA CGGGATCAGC 480

AGTACCAGGA TCTATTCACT TCGGATGTGG ATCTGTAGCA AGCTCAACAG CTATAGCCTC 540

CACTGCCAAA CTCCTTTTTT TCCTTCCTAG CATTTACCAA GAGACGCTTT TTAGCCTGTT 600

ATCTGTGGGT GCTACTTTTT CCCCTATGGT CCC 660

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

GCCACGCGTT CAGTAAAAAA GACAGCCAAG 30

(2) INFORMATION FOR SEQ ID NO: 31:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: YES
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

GCCCAGCGTA ACTACAGGCA CTGTGAGG 28

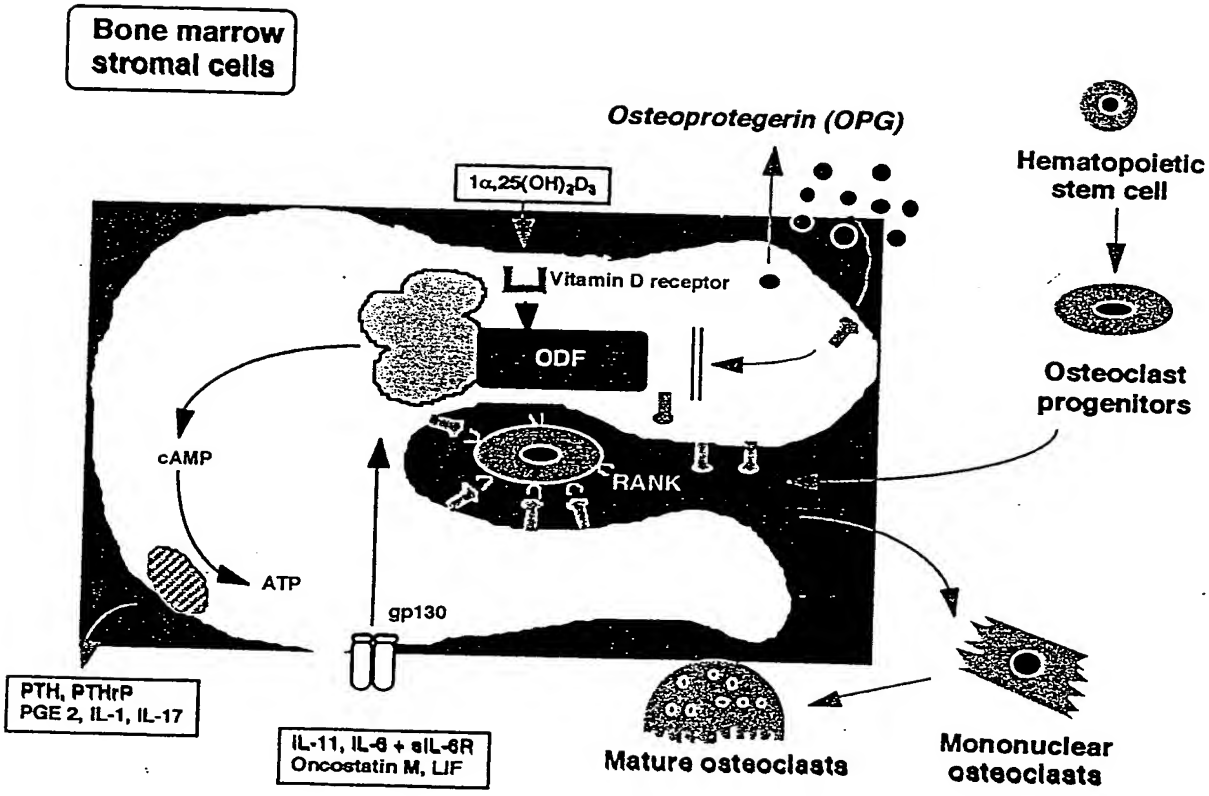

Bone marrow stromal cells
Osteoprotegerin (OPG)
Hematopoietic stem cell
1α,25(OH)2D3
Vitamin D receptor
ODF
Osteoclast progenitors
cAMP
RANK
ATP
gp130
PTH, PTHrP
PGE 2, IL-1, IL-17
IL-11, IL-6 + sIL-6R
Oncostatin M, LIF
Mature osteoclasts
Mononuclear osteoclasts

FIGURE 1

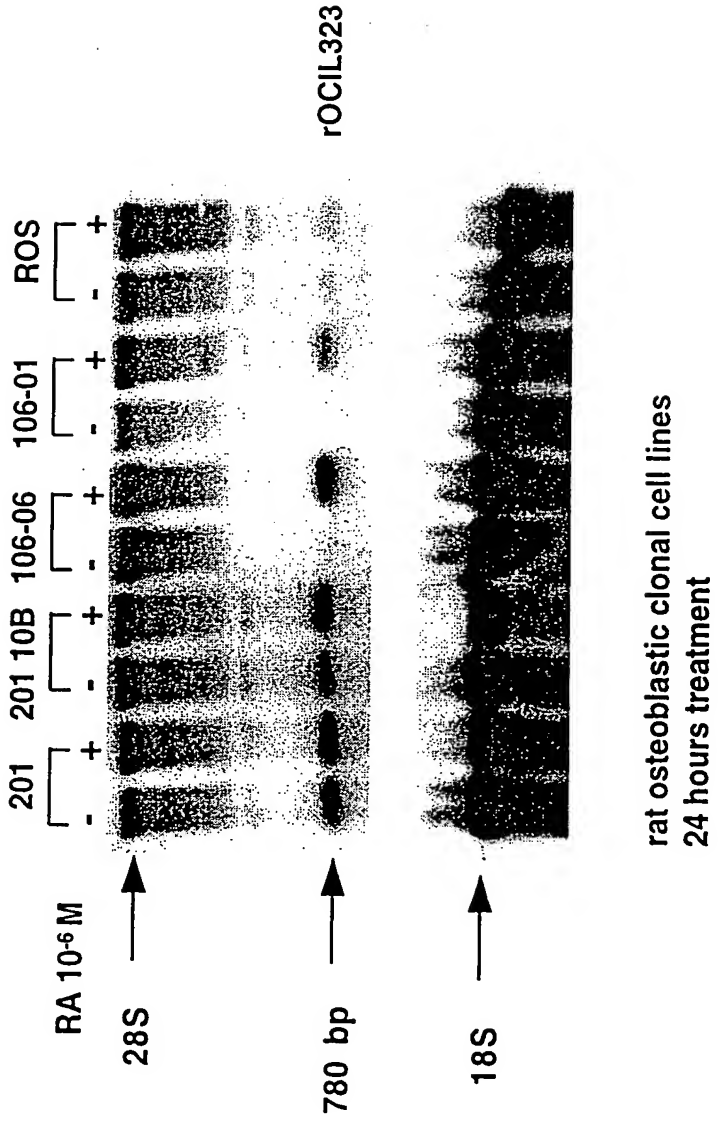

rOCIL323
ROS
106-01
106-06
201 10B
201
RA 10-6 M
28S
780 bp
18S
rat osteoblastic clonal cell lines
24 hours treatment

Figure 2

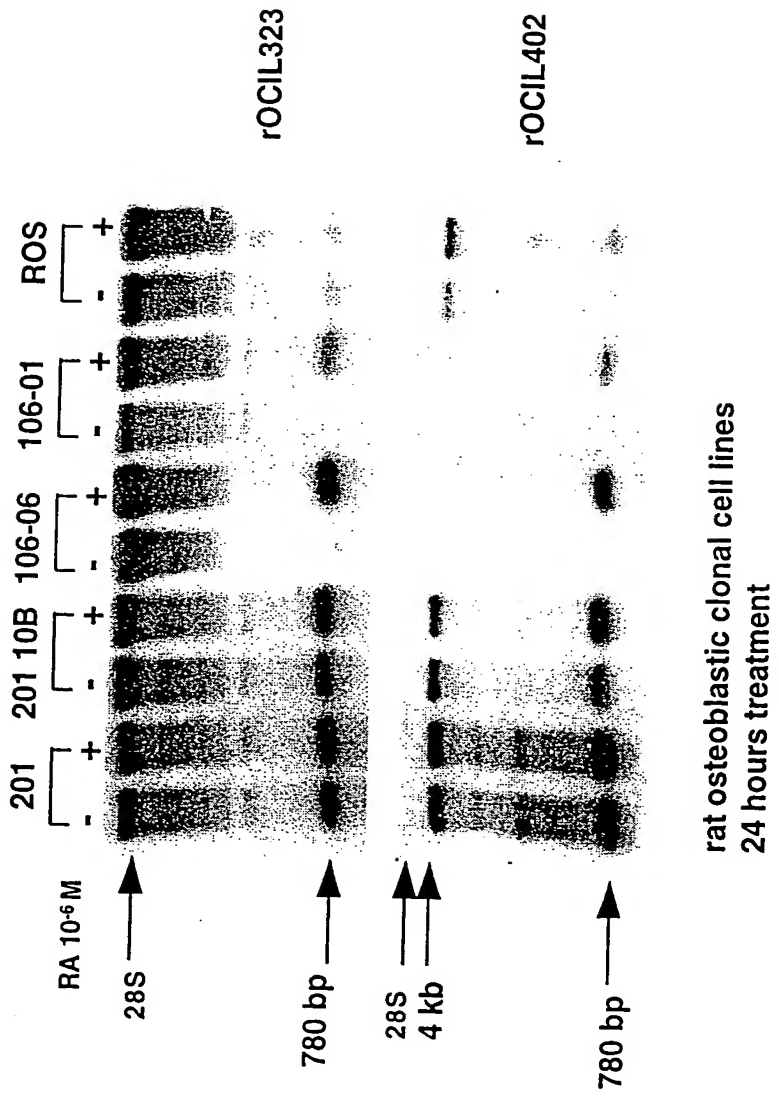

rOCIL323
rOCIL402
ROS
106-01
106-06
201 10B
201
RA 10-6 M
28S
780 bp
28S
4 kb
780 bp
rat osteoblastic clonal cell lines
24 hours treatment

Figure 3

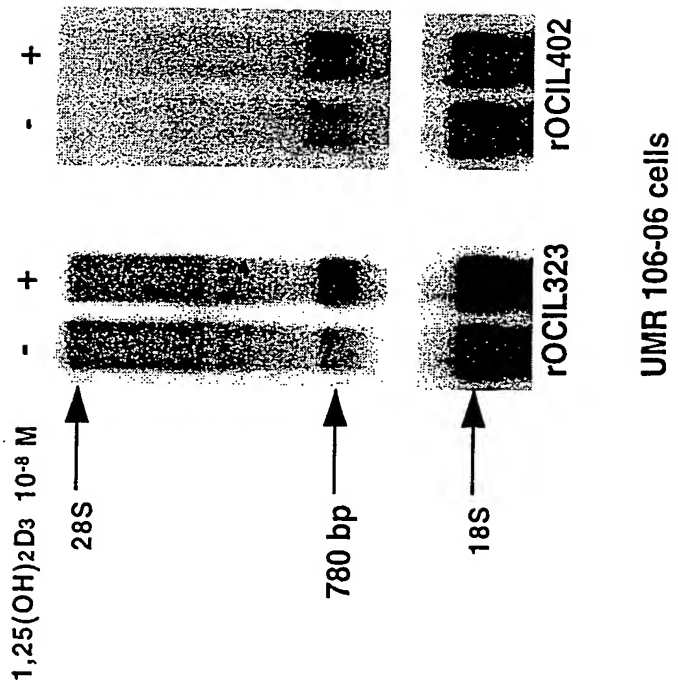

1,25(OH)2D3 10-8 M
-
+
-
+
28S
780 bp
18S
rOCIL323
rOCIL402
UMR 106-06 cells

Figure 4

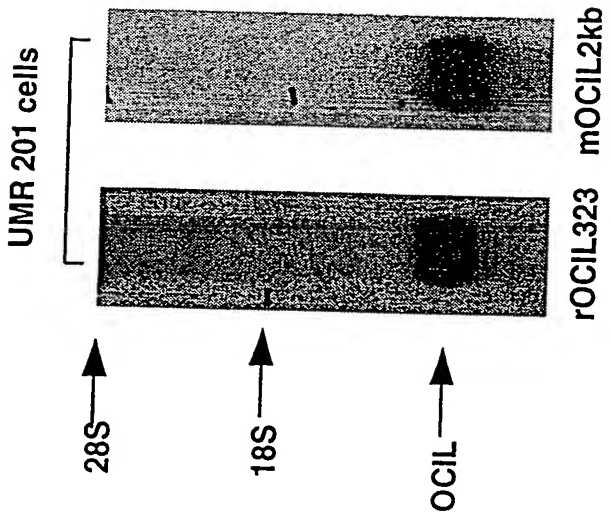
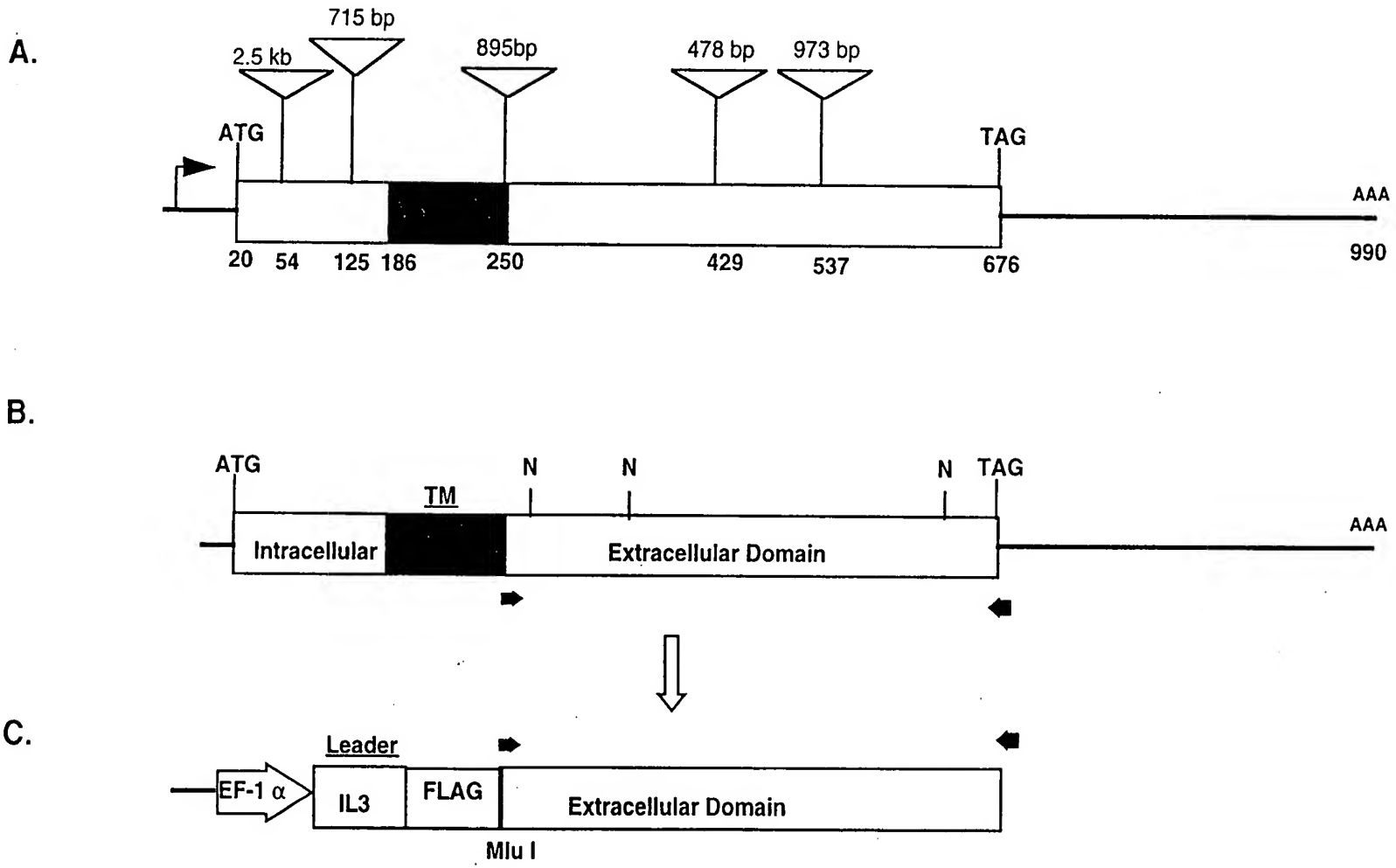


FIGURE 5

FIGURE 9



A: mOCIL Gene structure ▭ Open reading frame; ▮ Transmembrane domain; Y Intron

B: Predicted structure mOCIL. ▭ Open reading frame; ▮ Transmembrane domain; N Glycosylation sites.

C: Construct for recombinant protein expression. ➡ 5' PCR primer; ⬅ 3' PCR primer.

ClustalW Formatted Alignments

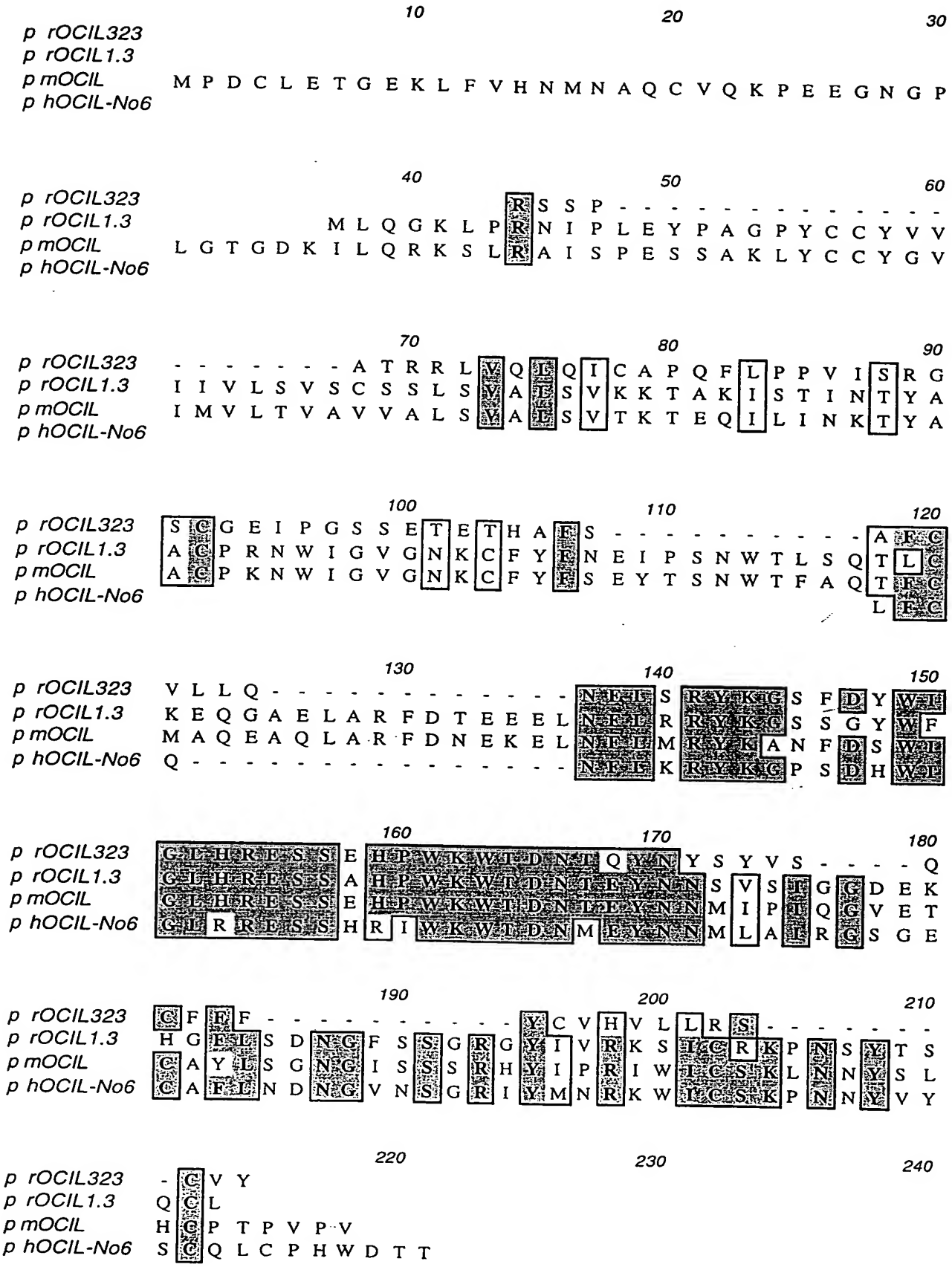


FIGURE 7

ClustalW Formatted Alignments

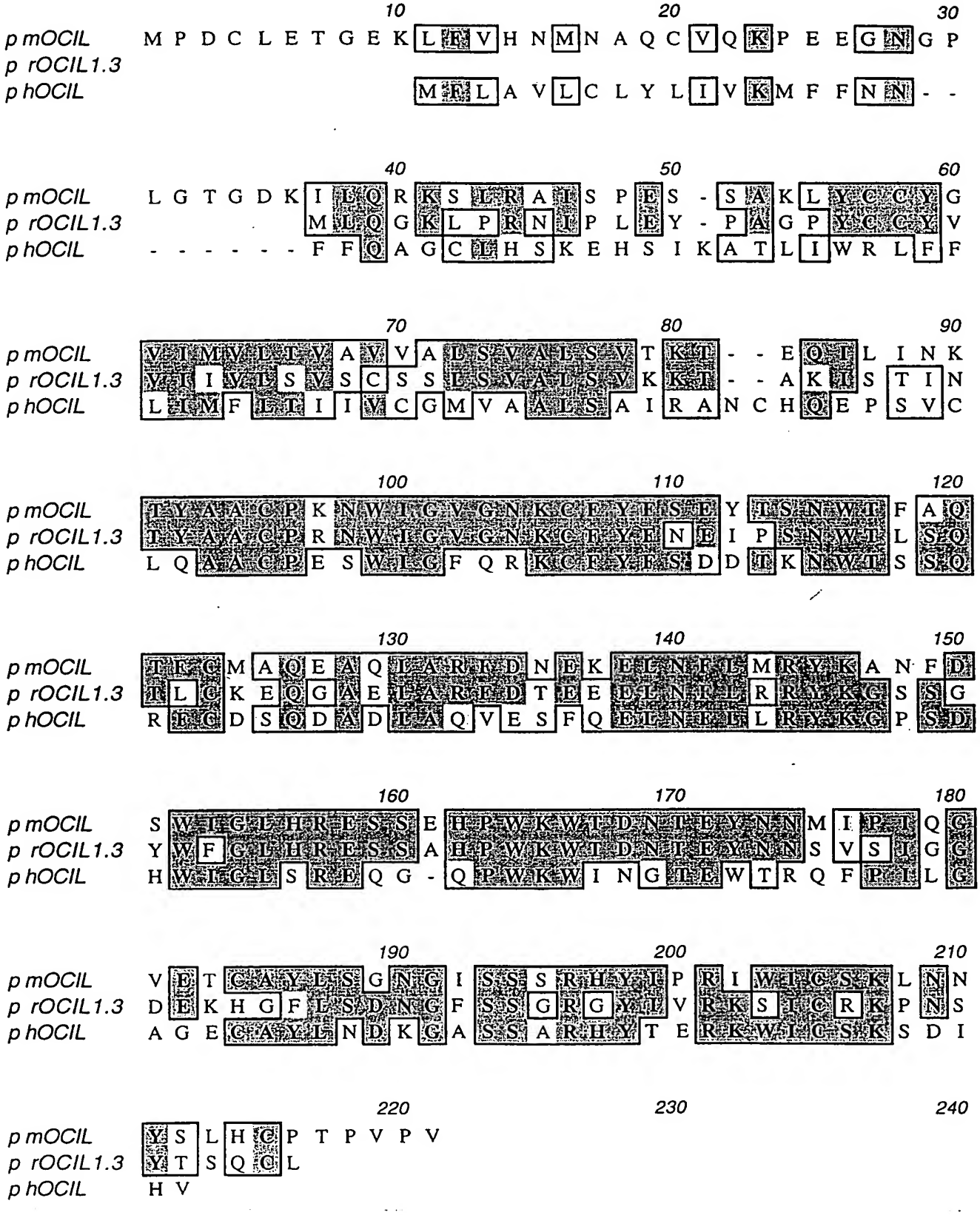
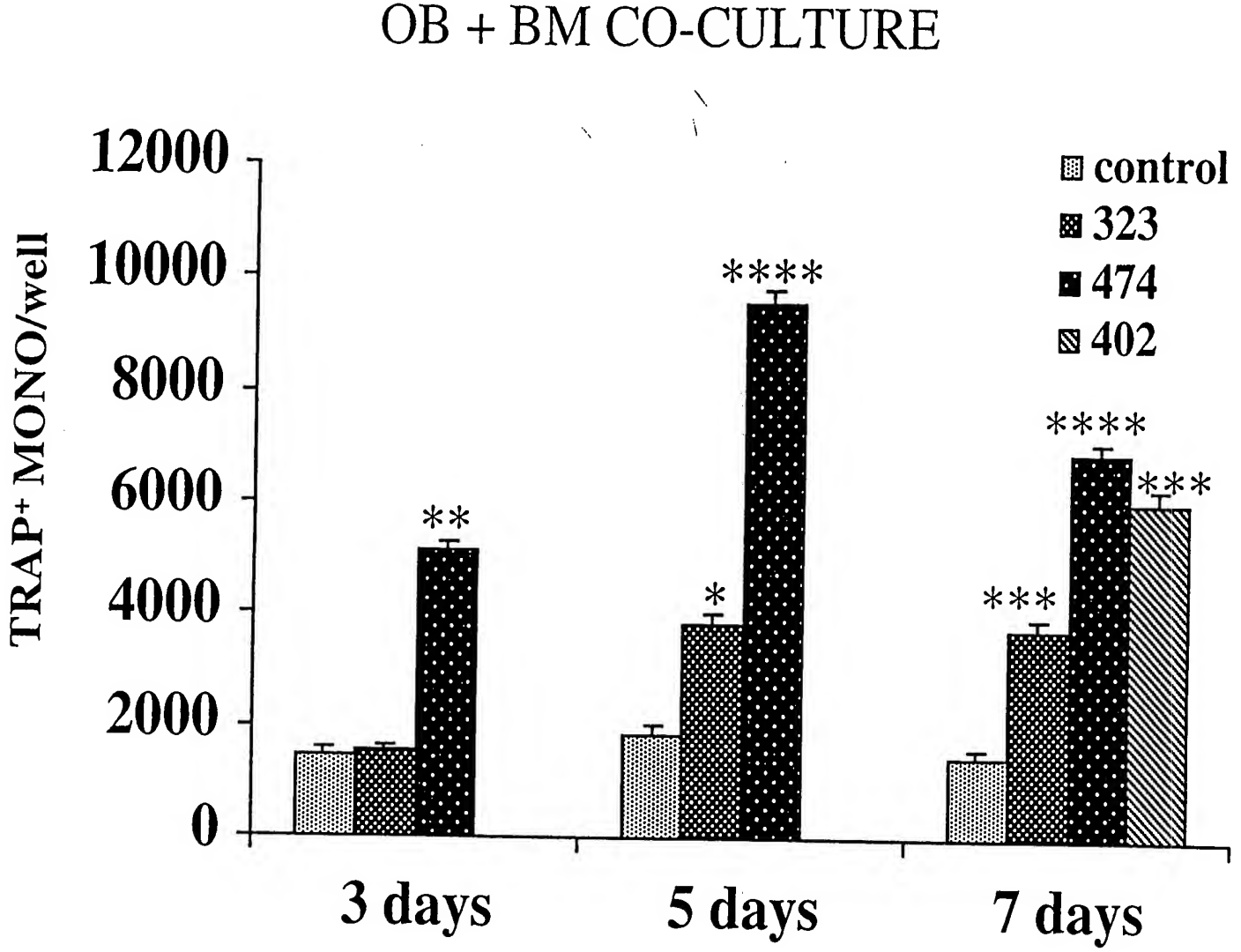


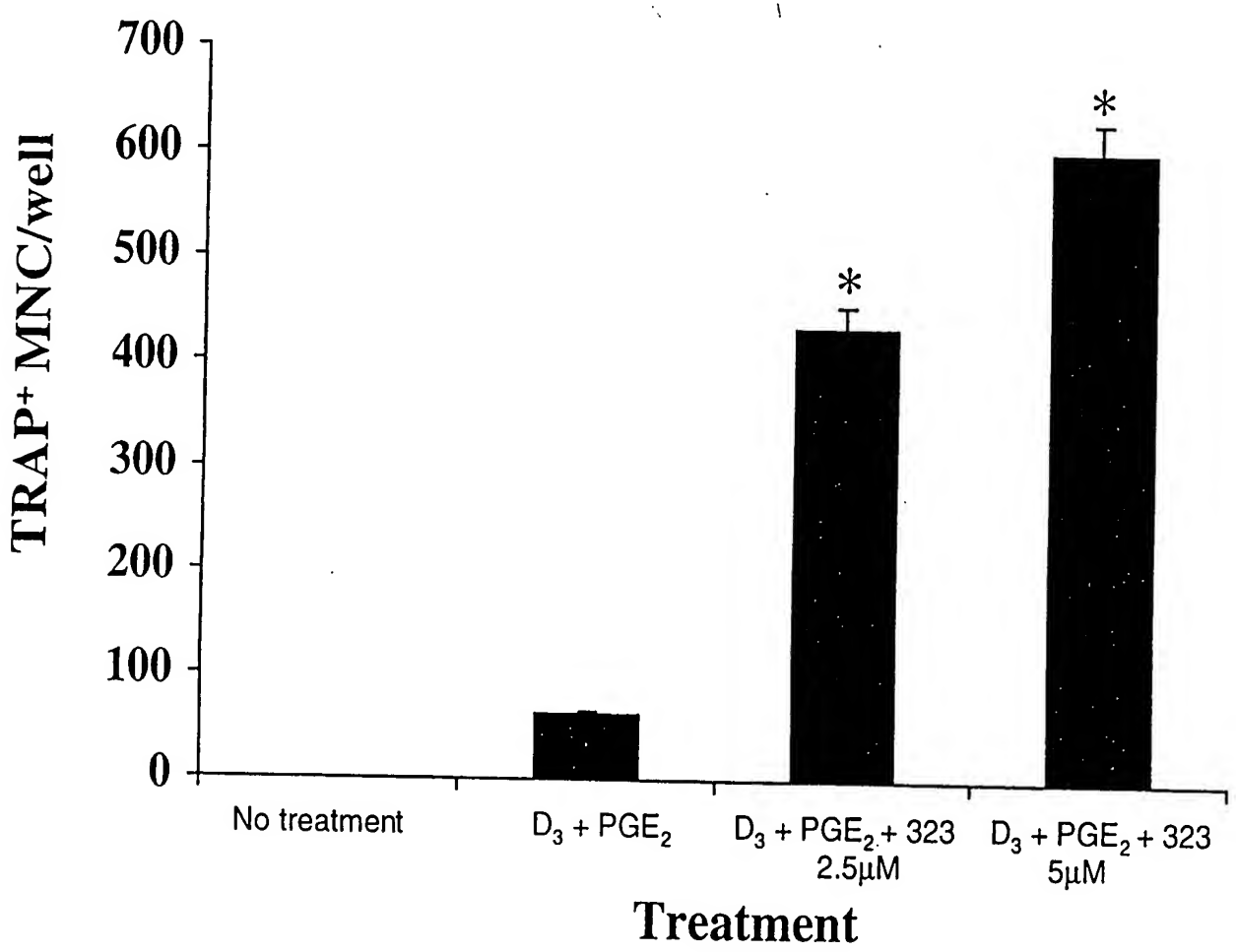
FIGURE 8



TRAP: tartrate-resistant acid phosphatase
MONO: mononucleate osteoclast precursors
*p < 0.025 vs control; **p < 0.005 vs control; ***p < 0.001 vs control; ****p < 0.0001 vs control

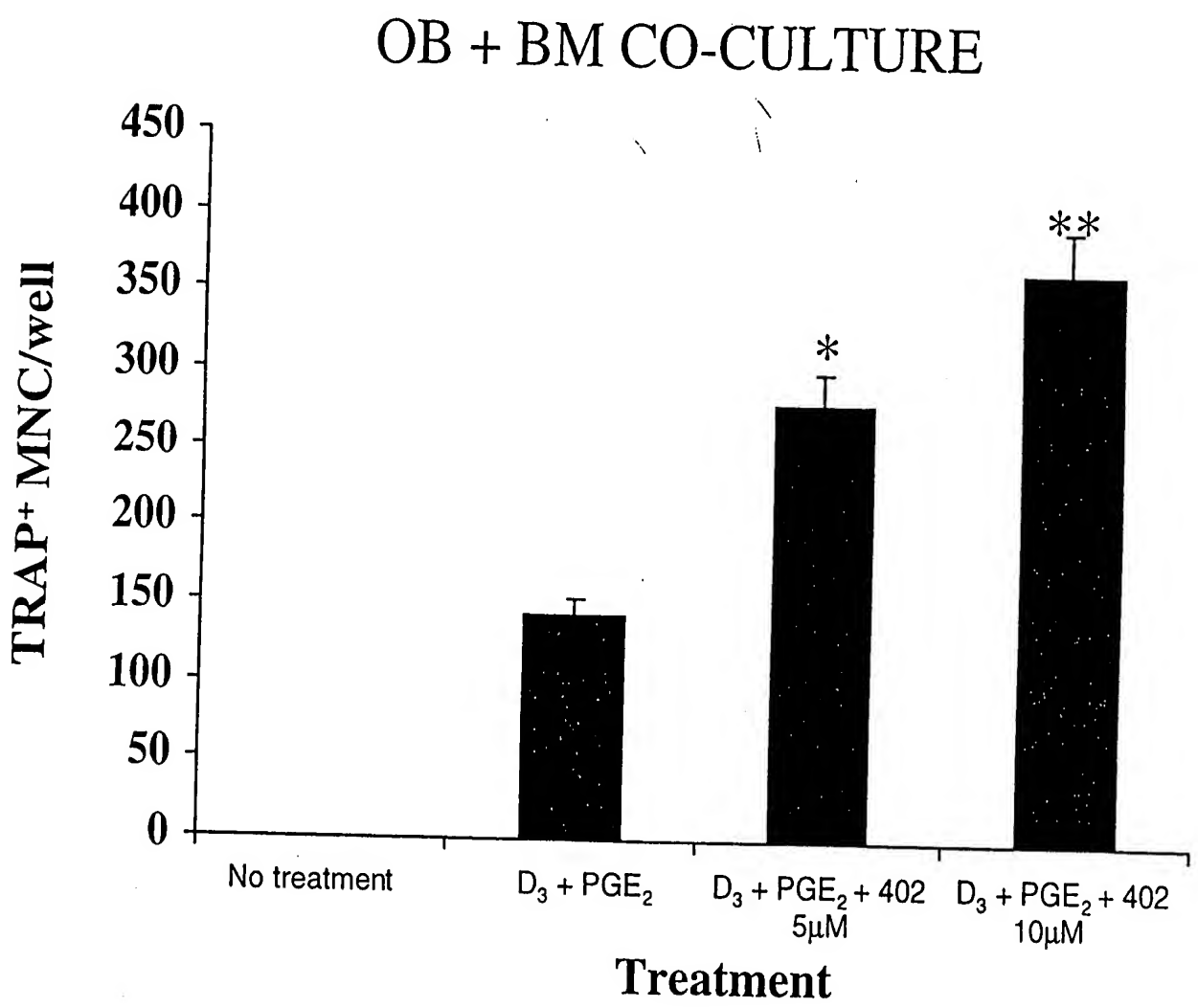
FIGURE 9a

OB + BM CO-CULTURE



TRAP: tartrate-resistant acid phosphatase
MNC: multinucleate osteoclasts
*p < 0.0001 vs D_3 + PGE_2 treated

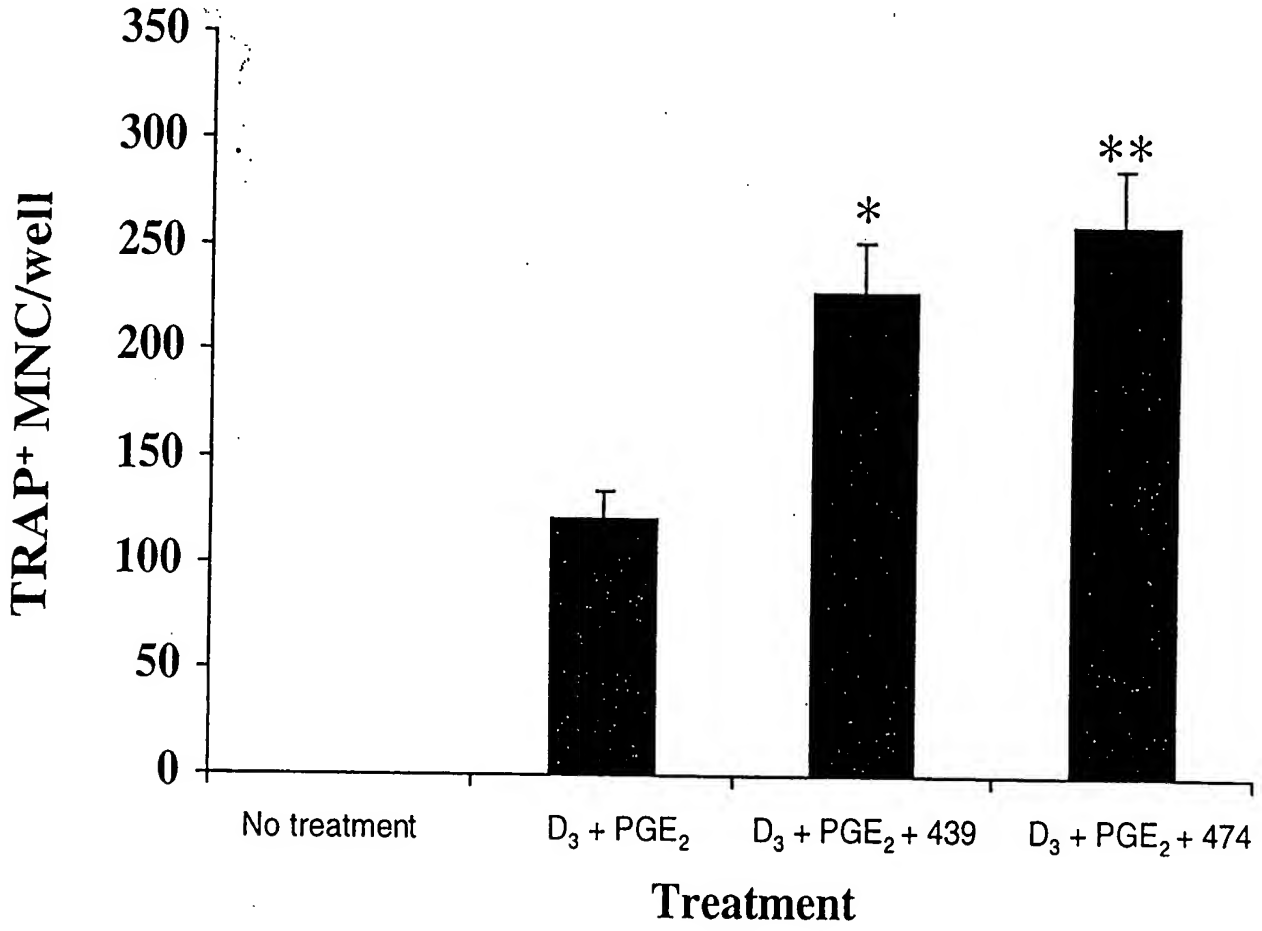
FIGURE 9b



TRAP: tartrate-resistant acid phosphatase
MNC: multinucleate osteoclasts
*$p < 0.01$ vs $D_3 + PGE_2$ treated; **$p < 0.0005$ vs $D_3 + PGE_2$ treated

FIGURE 9c

OB + BM CO-CULTURE



TRAP: tartrate-resistant acid phosphatase
MNC: multinucleate osteoclasts
*$p < 0.025$ vs $D_3 + PGE_2$ treated; **$p < 0.005$ vs $D_3 + PGE_2$ treated

FIGURE 9d

UMR 106 parental cells 24 hours treatment
OCIL
18S
control
TNF α 10ng/ml
DEX 10-7 M
PGE2 10-7 M
1,25(OH) 2D3 10-8 M
1,25(OH) 2D3 + PGE2
control
RA 10-6 M
control
PTH 100ng/ml

FIGURE 10

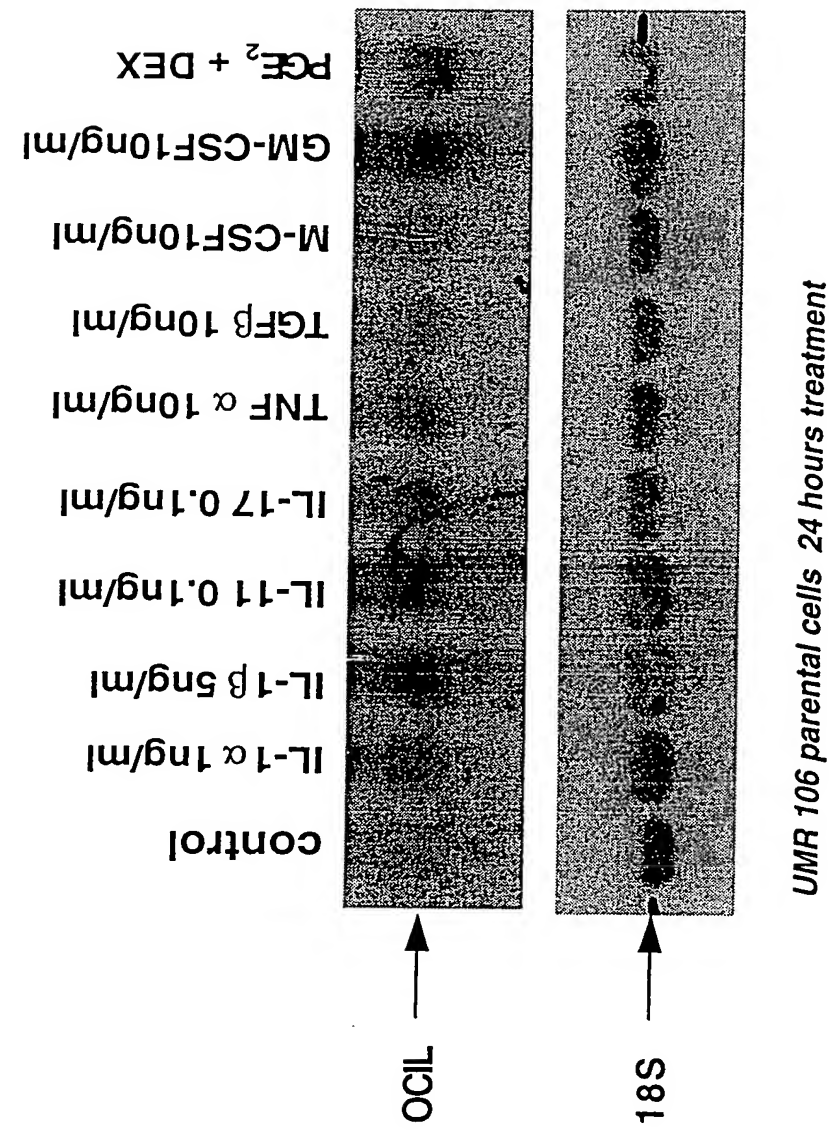


FIGURE 10 (cont.)

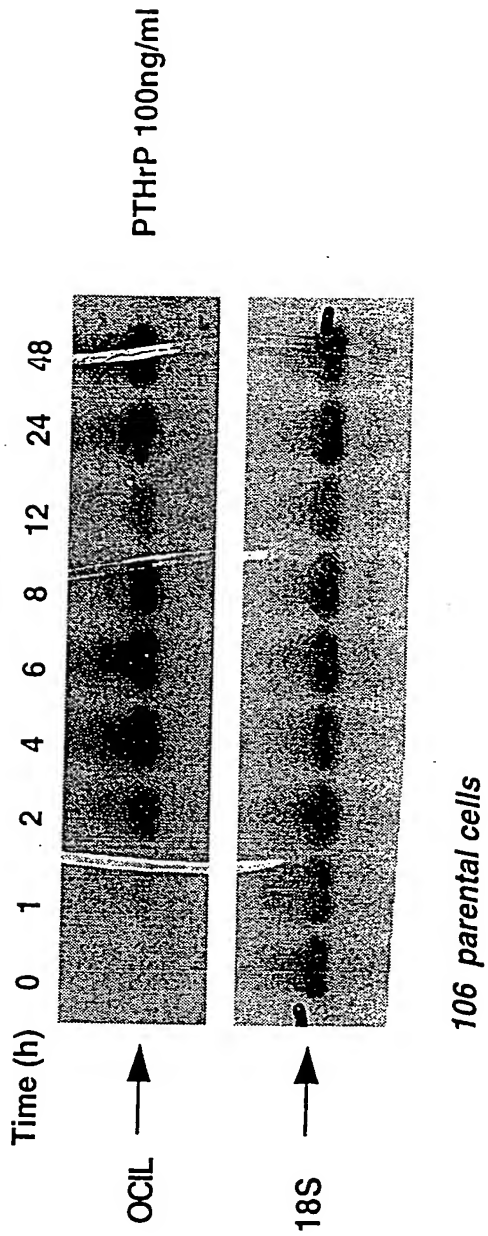

PTHrP 100ng/ml
Time (h) 0 1 2 4 6 8 12 24 48
OCIL
18S
106 parental cells

FIGURE 11A

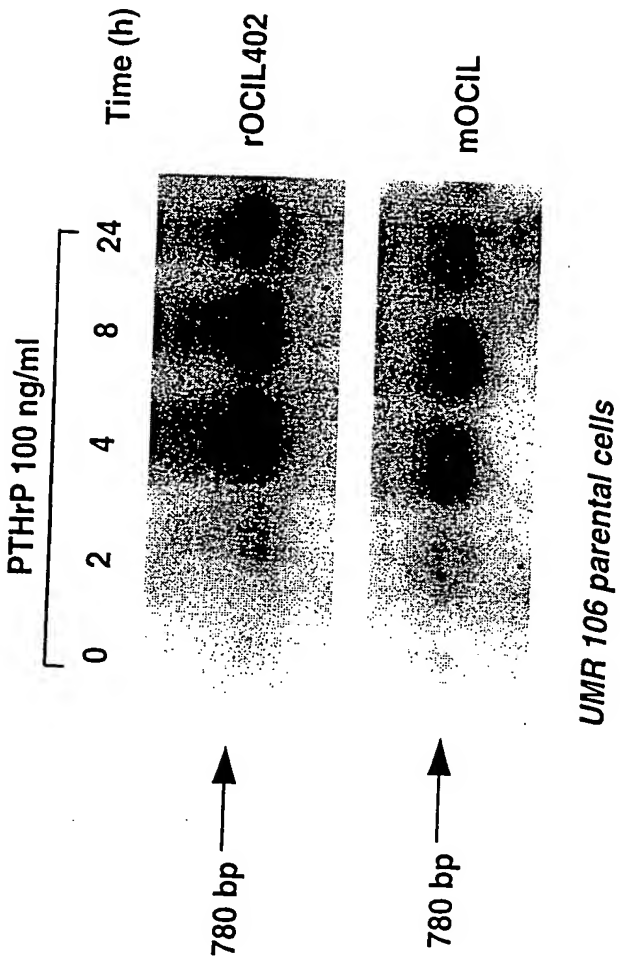

PTHrP 100 ng/ml
0
2
4
8
24
Time (h)
rOCIL402
mOCIL
UMR 106 parental cells
780 bp
780 bp

Figure 11B

Primary calvarial osteoblasts (POB)
24 hours treatment
18S
OCIL
control
IL-1α 1ng/ml
IL-1β 5ng/ml
IL-11 0.1ng/ml
DEX 10-7 M
PGE2 10-7 M
1,25(OH) 2D3 10-8 M
M-CSF 10ng/ml
RA 10-6 M

FIGURE 12

Mouse stromal cells (ST2)
OCIL
control
PGE2 10-7 M
DEX 10-7 M
IL-11 0.1ng/ml
1,25(OH) 2D3 10-8 M
1,25(OH) 2D3 + PGE2
1,25(OH) 2D3 + PGE2 + DEX

Figure 13A

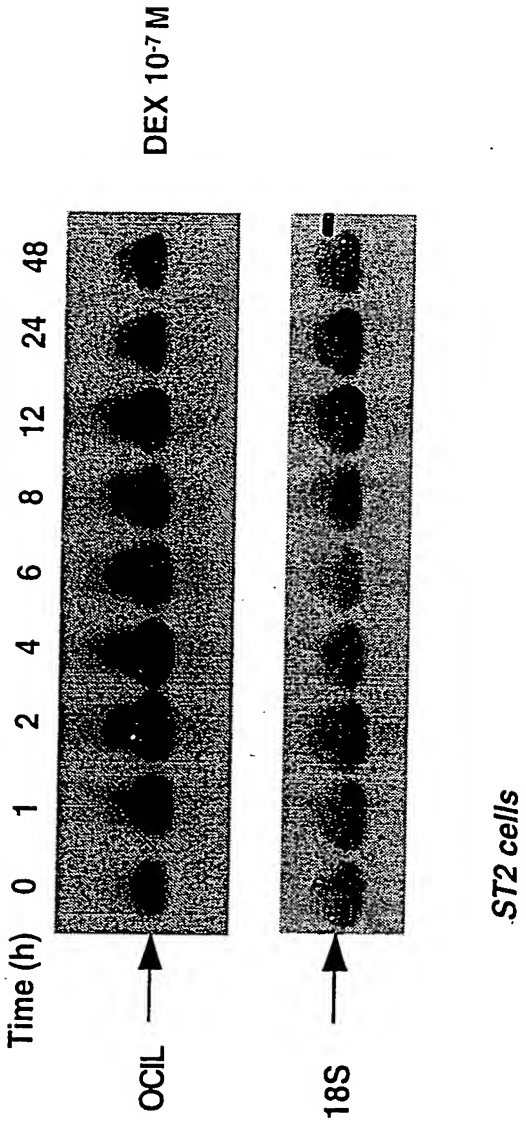


FIGURE 13B

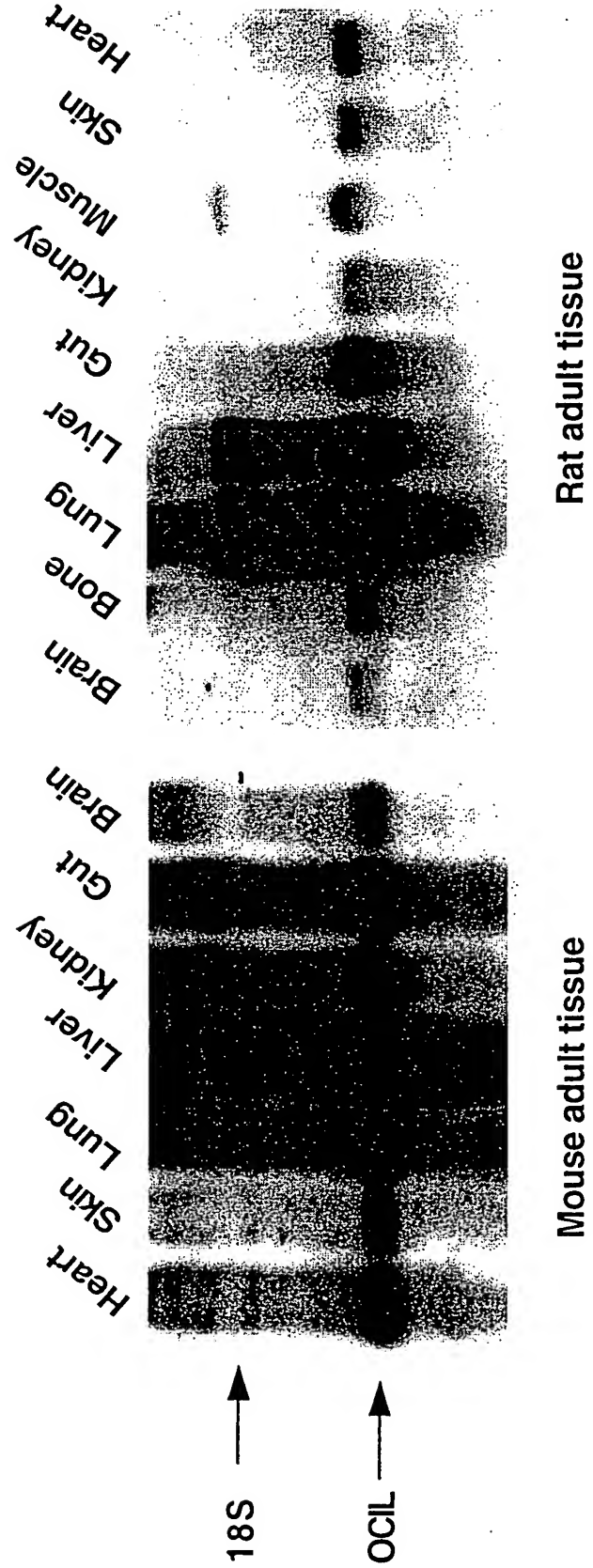
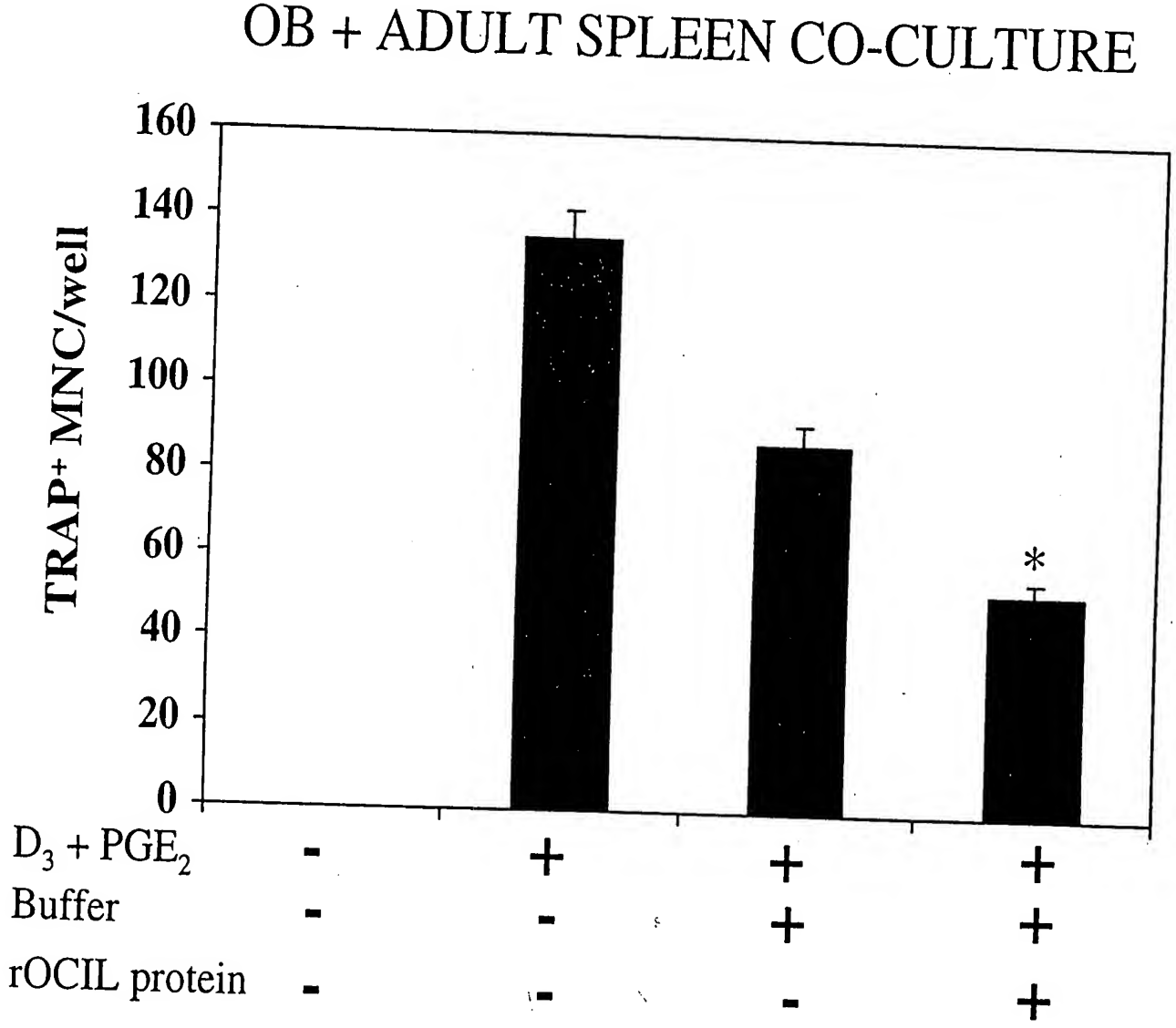

Heart
Skin
Muscle
Kidney
Gut
Liver
Lung
Bone
Brain
Rat adult tissue
Brain
Gut
Kidney
Liver
Lung
Skin
Heart
Mouse adult tissue
18S
OCIL

Figure 14

Figure 15



TRAP: tartrate-resistant acid phosphatase
MNC: multinucleate osteoclasts
*$p < 0.025$ vs $D_3 + PGE_2$ treated with buffer alone